ALMOST OPTIMAL LOCAL WELL-POSEDNESS OF THE MAXWELL-KLEIN-GORDON EQUATIONS IN $1+4$ DIMENSIONS

Sigmund Selberg*
Department of Mathematics
Johns Hopkins University
Baltimore, MD 21218

Abstract

We prove that the Maxwell-Klein-Gordon system on $\mathbb{R}^{1+4}$ relative to the Coulomb gauge is locally well-posed for initial data in $H^{1+\varepsilon}$ for all $\varepsilon > 0$. This builds on previous work by Klainerman and Machedon [6] who proved the corresponding result, with the additional restriction of small-norm data, for a model problem obtained by ignoring the elliptic features of the system, as well as cubic terms.

1 Introduction

The purpose of this paper is to prove local well-posedness (LWP) of the Maxwell-Klein-Gordon (MKG) equations on $\mathbb{R}^{1+4}$, relative to the Coulomb gauge, for initial data in $H^{1+\varepsilon}$, any $\varepsilon > 0$. This result is optimal in the sense that the critical Sobolev exponent for MKG on $\mathbb{R}^{1+4}$ is $s_c = 1$, and one does not expect well-posedness in H^s for s below this critical value; see the introduction in [8] and section 1.3 below, where we also make some remarks on the open question of well-posedness in the critical data norm H^1.

The analogous result for a hyperbolic model problem, obtained from the MKG system (6) below by setting the non-dynamical variable $A_0 \equiv 0$ and ignoring all cubic terms, was proved by Klainerman-Machedon [6], for small-norm initial data. That result was reproved, using different norms, and without any smallness assumption on the data, in the recent survey article [7]. The proof given there also used some ideas from [8], where the corresponding model problem for the Yang-Mills equation is considered.

*Current address: Inst. f. Mathematik, Univ. Wien, Strudlhofgasse 4, A-1090 Wien, Austria

The present work builds on the treatment of the model problem in [7]: To obtain *a priori* estimates on solutions of MKG with the requisite regularity, we complement the bilinear estimates proved there with estimates for cubic terms, and terms involving the non-dynamical variable, which satisfies an elliptic equation. It should be emphasized that the difficulty is to obtain LWP when s is very close to $s_c = 1$. If s is sufficiently large, one can prove LWP by much simpler methods than those employed here. See section 1.3 and Remark 1 in section 1.4.

Our method here can be modified[1] to treat the full Yang-Mills system in $\mathbb{R}^{1+4}$, proving LWP in $H^{1+\varepsilon}$, but only for initial data with small norm. This extends the result of Klainerman-Tataru [8] on a model equation for Yang-Mills. The reason for the small-norm restriction is that the elliptic equation in the Yang-Mills system relative to the Coulomb gauge is far more complicated than the one for MKG, and not in general globally solvable. To avoid this problem one can include the elliptic variable in the Picard iteration. Then to close the iteration one must assume small-norm data, since there is no way of compensating for large data by letting the existence time go to zero, as one can for an iteration involving only hyperbolic equations in a subcritical regime. Of course, using Picard iteration for an elliptic equation seems somewhat contrived. A better approach for Yang-Mills on $\mathbb{R}^{1+4}$ may be to work in the temporal gauge, as Tao [17] has successfully done for the case of $\mathbb{R}^{1+3}$. We hope to address this in a future paper.

Most of the previous work on MKG has been in dimension $1 + 3$. Let us summarize the known results for this case. LWP in the energy norm H^1 was proved by Klainerman and Machedon [4]. By conservation of the MKG energy, their result implies global well-posedness. In particular, they recovered an earlier global regularity result of Eardley and Moncrief [2] for smooth data. Cuccagna [1] proved LWP for small-norm data in H^s, $s > 3/4$. For $1 + 3$ dimensions, the critical regularity is $s_c = 1/2$, but the question of LWP below $s = 3/4$ remains open. In both [4] and [1] the Coulomb gauge is used. More recently, Tao [17] has proved small-norm LWP for $s > 3/4$ using the temporal gauge, for the more general Yang-Mills equations.

Our method here can be used to remove the small-norm restriction in the result of Cuccagna. The essential reason for this limitation in [1] is that the elliptic variable was included in the iteration. If instead one solves the elliptic equation and reduces to a purely hyperbolic system, as we do here, this obstruction is removed, and one can get a large data LWP result. A crucial fact needed to make this work is that in the Klainerman-Machedon bilinear estimates used by Cuccagna, the space-time derivative $|D_{t,x}|^{-a}$ acting on the product can be replaced by $|D_x|^{-a}$, as observed in [11] (cf. also the remark in the Appendix), rendering unnecessary the decomposition in Fourier space used in [1]. See also Remark 3 in section 4.

[1] We do not prove this here, but hope to address it in a separate paper dealing with the Yang-Mills equations on $\mathbb{R}^{1+4}$ in Coulomb as well as temporal gauge. Note that Yang-Mills essentially contains MKG as a special case.

Finally, we remark that our proof should generalize without difficulty to the higher dimensional case of MKG on $\mathbb{R}^{1+n}$ with $n \geq 5$, giving LWP for $s > s_c = \frac{n-2}{2}$. In fact, the difficulty of the problem decreases with increasing dimension.

1.1 The Maxwell-Klein-Gordon system

The Klein-Gordon equation can be derived as a relativistic analogue of the Schrödinger equation for a free particle. It is obtained from the relativistic energy-momentum relation $E^2 = \mathbf{p}^2c^2 + m^2c^4$, where E is the energy of the particle, $m > 0$ its rest mass, $\mathbf{p}$ its momentum and c the light speed. Setting $c = 1$ from now on, and applying the quantum mechanical principle of replacing classical quantities by operators:

Energy $\qquad E \longrightarrow i\frac{\partial}{\partial t}$,

Momentum $\qquad \mathbf{p} \longrightarrow \frac{1}{i}\nabla$,

one obtains the free Klein-Gordon equation

$$\Box\phi = m^2\phi, \tag{1}$$

where $\phi(t,x) \in \mathbb{C}$ and $\Box = \partial_\mu\partial^\mu = -\partial_t^2 + \Delta$ is the wave operator on $\mathbb{R}^{1+n}$. Here we use relativistic coordinates $t = x^0, x^1, \ldots, x^n$ on the Minkowski spacetime $\mathbb{R}^{1+n}$ with the metric $\mathrm{diag}(-1, 1, \ldots, 1)$; indices are raised and lowered relative to this metric, and the Einstein summation convention is in effect: roman indices $j, k, \ldots$ run from 1 to n, greek indices $\mu, \nu, \ldots$ from 0 to n. We write ∂_μ for $\frac{\partial}{\partial x^\mu}$, and $\partial_t = \partial_0$. We shall use $\Re z$ and $\Im z$ to denote the real and imaginary parts of $z \in \mathbb{C}$.

The coupling of (1) to an electromagnetic field represented by a potential $A_\mu(t,x) \in \mathbb{R}$ is achieved by the so-called minimal substitution

$$\partial_\mu \longrightarrow D_\mu = \partial_\mu + iA_\mu,$$

where iA_μ acts as a multiplication operator. This gives

$$D_\mu D^\mu\phi = m^2\phi. \tag{2}$$

which is the Klein-Gordon equation. It has an associated current density

$$j_\mu = \Im\left(\phi\overline{D_\mu\phi}\right) = \Im\left(\phi\overline{\partial_\mu\phi}\right) - A_\mu\left|\phi\right|^2, \tag{3}$$

satisfying the conservation law

$$\partial^\mu j_\mu = 0. \tag{4}$$

In fact, one has the general identity $\partial_\mu\Im\left(\phi\overline{D^\mu\phi}\right) = \Im\left(\phi\overline{D_\mu D^\mu\phi}\right)$, so (4) follows immediately from (2).

The Maxwell-Klein-Gordon system is then obtained by coupling (2) to the Maxwell equation

$$\partial^\nu F_{\mu\nu} = j_\mu \tag{5}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the electromagnetic field tensor and j_μ is the Klein-Gordon current density (3). The system (5),(3),(2) is then what we — provisionally — call the Maxwell-Klein-Gordon system. We want to consider this as a system of second order PDE in the unknowns A_μ and ϕ, but there is an obvious problem with this, since $F_{\mu\nu}$ — and hence the observables, i.e., the electric and magnetic field vectors, whose components are entries of the matrix $F_{\mu\nu}$ — are not uniquely determined by A_μ. This is known as the gauge ambiguity, and to resolve it one adds another equation to the system, a so-called gauge condition, which uniquely determines A_μ. The standard gauge conditions are (i) Lorentz: $\partial^\mu A_\mu = 0$, (ii) Coulomb: $\partial^i A_i = 0$ and (iii) temporal: $A_0 = 0$.

In this paper, we shall rely on the Coulomb condition, which carries the advantage — as Klainerman and Machedon observed in [4] for the case of $n = 3$ — that the bilinear terms involving derivatives turn out to be of null form type, and therefore have better regularity properties than generic products. Since the derivation of the null form structure in [4] uses the special vector calculus of $n = 3$, in particular the curl operator, we include a generalization of this argument to arbitrary dimension in section 1.5.

1.2 Main result

If we add the Coulomb gauge condition $\partial^j A_j = 0$ to the MKG system (5), (3), (2) and expand, we get:

$$\Delta A_0 = -\Im\big(\phi\overline{\partial_t\phi}\big) + |\phi|^2 A_0, \tag{6a}$$

$$\Box A_j = -\Im\big(\phi\overline{\partial_j\phi}\big) + |\phi|^2 A_j - \partial_j\partial_t A_0, \tag{6b}$$

$$\Box\phi = -2iA^j\partial_j\phi + 2iA_0\partial_t\phi + i(\partial_t A_0)\phi + A^\mu A_\mu\phi + m^2\phi, \tag{6c}$$

$$\partial^j A_j = 0. \tag{6d}$$

In the rest of the paper, with the exception of section 1.3, we will take $n = 4$. Thus, the unknowns are

$$A_0, A_j : \mathbb{R}^{1+4} \to \mathbb{R}, \quad \phi : \mathbb{R}^{1+4} \to \mathbb{C}.$$

When convenient, we shall write A for the four-vector field $(A^j)_{j=1,\dots,4}$. Initial data are specified at time $t = 0$:

$$A\big|_{t=0} = a \in H^s, \qquad \partial_t A\big|_{t=0} = b \in H^{s-1}, \tag{7a}$$

$$\phi\big|_{t=0} = \phi_0 \in H^s, \qquad \partial_t\phi\big|_{t=0} = \phi_1 \in H^{s-1}, \tag{7b}$$

where $H^s = \big\{ f \in \mathcal{S}'(\mathbb{R}^4) : (I-\Delta)^{s/2} f \in L^2(\mathbb{R}^4) \big\}$ and a, b are real vector fields. In view of the Coulomb condition (6d), we must require

$$\partial^j a_j = \partial^j b_j = 0. \tag{8}$$

Observe that no data are specified for the non-dynamical variable A_0. This is quite natural, because A_0 is determined by ϕ and $\partial_t \phi$ at any time t by solving the elliptic equation (6a).

Theorem 1. *For all $s > 1$, the Cauchy problem* (6),(7),(8) *on $\mathbb{R}^{1+4}$ is locally well-posed.*

Local well-posedness here includes (a) existence of a local solution

$$A_0 \in C\big([0,T], \dot{H}^1\big) \cap C^1\big([0,T], L^2\big) \tag{9a}$$

$$A_j, \phi \in C\big([0,T], H^s\big) \cap C^1\big([0,T], H^{s-1}\big) \tag{9b}$$

up to a time $T > 0$ depending continuously on the H^s-norm of the initial data; (b) uniqueness of the solution; (c) continuous dependence on the data; and (d) persistence of higher regularity. A more precise statement, for an equivalent system, can be found in Theorem 2, section 1.5. In particular, the uniqueness is proved not in the class (9), but in a smaller space determined by the iteration norms; see (20).

To prove Theorem 1 we shall in effect eliminate the nondynamical variable A_0 from the equations, by solving the elliptic equations. This leaves us with a system of nonlinear wave equations, which we then prove is locally well-posed. Once this has been achieved, we can go back to the original system (6), and conclude that this is also well-posed.

Let us be more precise. We introduce a new variable $B_0 = \partial_t A_0$. Applying ∂^j to (6b) and using (6d) yields

$$\Delta B_0 = -\Im \partial^j \left(\phi \overline{\partial_j \phi} \right) + \partial^j \left(|\phi|^2 A_j \right). \tag{10}$$

Now we eliminate A_0 and $\partial_t A_0 = B_0$ from (6b) and (6c) by solving (6a) and (10). Thus $A_0 = A_0(\phi)$ and $B_0 = B_0(A, \phi)$ are nonlinear operators. Since the Coulomb condition (6d) turns out to be automatically satisfied because of the constraint (8), we obtain a system of nonlinear wave equations

$$\Box A = \mathcal{M}(A, \phi), \tag{11a}$$

$$\Box \phi = \mathcal{N}(A, \phi), \tag{11b}$$

where $\mathcal{M}$ and $\mathcal{N}$ are certain operators[2], nonlocal in the space variable, which are sums of terms of the following types: (i) bilinear and higher order multilinear expressions involving A and ϕ and their first derivatives, (ii) terms involving $A_0(\phi)$, and (iii) a linear term $m^2\phi$ in (11b). Moreover, all the bilinear terms have a null structure, due to the Coulomb gauge, and for these terms one already has good estimates (see [6], and also [8] for the case of Yang-Mills; here we shall rely more particularly on variants of these estimates proved in [7]). We complement these with estimates for the higher order multilinear terms and terms containing $A_0(\phi)$, and local well-posedness of the system (11) then follows by the general theory developed in the author's paper [12].

[2]See section 1.5 for precise definitions

Then the original system (6) is also locally well-posed, by reversing the steps leading to (11). That is, if (A, ϕ) has the requisite regularity (see (20)) and solves (11) on a time-slab, and if we set $A_0 = A_0(\phi)$, then $\partial_t A_0 = B_0(A, \phi)$ in the sense of distributions and the triple (A_0, A, ϕ) solves (6) on the same time-slab.

Thus, we show that the systems (6) and (11) are *equivalent* for sufficiently regular solutions.

1.3 Scaling, optimality and the null condition

As for many other field theories, there are two types of "critical" behaviour associated to the MKG system on $\mathbb{R}^{1+n}$. On the one hand, there is the critical regularity s_c such that the homogeneous initial data space $\dot{H}^{s_c}$ is left invariant under the natural scaling transformation associated to MKG:

$$A_\mu(t,x), \phi(t,x) \longrightarrow \lambda A_\mu(\lambda t, \lambda x), \lambda \phi(\lambda t, \lambda x), \tag{12}$$

where λ is a positive parameter.[3] Since

$$\|\lambda f(\lambda x)\|_{\dot{H}^s} = \lambda^{s-(n-2)/2} \|f\|_{\dot{H}^s}, \tag{13}$$

we conclude that $s_c = \frac{n-2}{2}$. In general[4] one expects field theories to be locally well-posed (LWP) for $s > s_c$ and ill-posed for $s < s_c$; we say more about this below. In the critical case $s = s_c$ one expects some type of weakened well-posedness[5] for data with small norm.

On the other hand, there is the energy-critical dimension n such that the critical regularity is at the level of the energy:[6] $s_c = 1$. For MKG this means $n = 4$, which is the dimension we consider in this paper. For field theories in general, one expects global regularity in the critical dimension, as well as in subcritical dimensions ($s_c < 1$), and breakdown of regularity for large data in supercritical dimensions ($s_c > 1$).

As mentioned above, the global regularity is known in the subcritical dimension $n = 3$ for MKG, but the question of global regularity in the critical dimension $n = 4$, even for data with small energy, remains open. By conservation of energy, a LWP result, for small-norm data, at the critical regularity $s_c = 1$ would settle this question in the affirmative, but it is perhaps more realistic to expect a more direct proof of global regularity in analogy with the results of Tao [15, 16] for wave maps into a sphere. It is to be hoped that our almost optimal LWP result will play some role in any such result.

The expectation of ill-posedness for $s < s_c$ is based on the scaling (12) and (13). First, if blow-up occurs for smooth, compactly supported data, then one

[3] By this we mean that if A_μ, ϕ solve MKG, then so do the rescaled fields, although the rest mass changes from m to λm.

[4] See [7, Section 1.3] for further discussion and references.

[5] For example, one does not expect smooth dependence on initial data, which rules out proof by iteration. A good example is wave maps into a sphere; see Tao [15, Section 1] for a summary of the regularity results for wave maps.

[6] MKG has a conserved energy which is at the level of the H^1 data norm; see [4].

can construct data in H^s, $s < s_c$, with arbitrarily small norm, for which there is no local existence; see, e.g., [13, pp 98–99] for this argument. However, this is not a very convincing point to make here, as we do expect global regularity for MKG on $\mathbb{R}^{1+4}$. We can show, however, that it is impossible to prove any well-posedness result for $s < s_c$ using an iteration argument based on estimates. The idea can be illustrated by the following example: As is well-known, the algebra inequality

$$\|fg\|_{H^s} \leq C_{s,n} \|f\|_{H^s} \|g\|_{H^s} \tag{14}$$

holds for $H^s(\mathbb{R}^n)$ iff $s > n/2$. A rather crude way of ruling out the range $s < n/2$ is to observe that if (14) holds, then by rescaling[7] $x \to \lambda x$ and letting $\lambda \to \infty$, we get $1 \lesssim \lambda^{s-n/2}$.

This idea is easily applied to the iteration for MKG written in the form (11). Let us take $m = 0$ here to make the system scale invariant. If $a = b = 0$ in (7a) and $\phi_1 = 0$ in (7b), then the first iterate of A solves[8] $\Box A^{(1)} = -\mathcal{P}\Im(\phi^{(0)}\overline{\nabla\phi^{(0)}})$ with zero data, where $\phi^{(0)}$ is the solution of $\Box\phi^{(0)} = 0$ with data $(\phi_0, 0)$. If we can prove LWP in H^s by iteration, there must be an estimate

$$\sup_{0 \leq t \leq 1} \left\| A^{(1)}(t) \right\|_{H^s} \lesssim \|\phi_0\|_{H^s}^2 , \tag{15}$$

for all ϕ_0 with sufficiently small norm. Now assume $s < s_c$. We then claim that (15) implies $A^{(1)} \equiv 0$, which is absurd. Indeed, given $T > 0$, apply (15) to the rescaled iterate

$$\widetilde{A}^{(1)}(t,x) = \lambda A^{(1)}(\lambda t, \lambda x)$$

at time $t = T/\lambda$. As $\lambda \to \infty$ this gives

$$\lambda^{s-s_c} \left\| A^{(1)}(T) \right\|_{H^s} \lesssim (\lambda^{s-s_c})^2 \|\phi_0\|_{H^s}^2 ,$$

whence $A^{(1)}(T) = 0$.

Remark. This argument has nothing to do with the null condition, of course. A more careful analysis (see [7, Section 1]) suggests that for a generic equation of the form

$$\Box u = u\partial u$$

on $\mathbb{R}^{1+n}$ one needs $s \geq \max(\frac{n-2}{2}, \frac{n+1}{4})$ in order for the iterates to stay in H^s, and this is consistent with Lindblad's counterexamples [10]. However, if the right hand side is replaced by a null form expression like (23) or (24), then one only needs $s \geq \max(\frac{n-2}{2}, \frac{n-1}{4})$, so the null condition improves matters when $n \leq 4$.

As remarked already, the main difficulty is to prove LWP when s is very close to s_c, whereas simpler arguments can be used for larger s. Let us be more

[7]In the limit $\lambda \to \infty$, the inhomogeneous Sobolev norm H^s scales like $\dot{H}^s$.

[8]Here $\mathcal{P}$ denotes the projection onto divergence free vector fields. See section 1.5.

precise. Observe that relative to Lorentz gauge, MKG on $\mathbb{R}^{1+n}$ is a system of nonlinear wave equations of the schematic form (see [7, Section 1])

$$\square u = u\partial u + u^3, \tag{16}$$

and for this system LWP for $s > n/2$ can be proved by standard methods, just using the energy inequality for the wave equation and Sobolev embeddings. This can easily be improved to $s > \frac{n-1}{2}$ by using a $L_t^2 L_x^\infty$ spacetime estimate instead of just Sobolev embedding. For $n = 4$ this gives LWP for $s > 3/2$, which is still one quarter of a derivative above what one expects (cf. remark above) from the analysis of the first iterate of (16), namely $s > 5/4$. No proof of LWP of (16) in this range seems to exist in the literature, but it should be obtainable using the spaces $H^{s,\theta}$ (see section 1.4) and L^2 bilinear estimates for the homogeneous wave equation of the type first proved in [5]. However, to go below the regularity 5/4, one really needs the null condition, which seems to rule out Lorentz gauge. Of course, once a LWP result has been proved in one gauge, one can in principle use gauge transformations (see [4]) to transfer this result to other gauges; but to make this rigorous requires sufficient regularity of the solutions, and we will not consider this question here.

1.4 Function spaces

Here we define the spaces that we make use of. See [7] for more details.

The Fourier transform of $f(x)$ [resp. $u(t,x)$] is denoted $\widehat{f}(\xi) = \mathcal{F}f(\xi)$ [resp. $\widehat{u}(\tau,\xi) = \mathcal{F}u(\tau,\xi)$].

We say that a norm $\|\cdot\|$, on some space $\mathcal{X}$ of tempered distributions, depends only on the size of the Fourier transform if

$$|\widehat{u}| \le |\widehat{v}| \implies \|u\| \le \|v\| .$$

(Here we assume, of course, that the Fourier transform of any element of $\mathcal{X}$ is a function.)

If $\mathcal{X}$ and $\mathcal{Y}$ are two normed spaces, the notation $\mathcal{X} \hookrightarrow \mathcal{Y}$ means continuous inclusion.

For any $\alpha \in \mathbb{R}$ we define Fourier multiplier operators Λ^α, Λ_+^α and Λ_-^α by

$$\begin{aligned}
\widehat{\Lambda^\alpha f}(\xi) &= \left(1 + |\xi|^2\right)^{\alpha/2} \widehat{f}(\xi), \\
\widehat{\Lambda_+^\alpha u}(\tau,\xi) &= \left(1 + \tau^2 + |\xi|^2\right)^{\alpha/2} \widehat{u}(\tau,\xi), \\
\widehat{\Lambda_-^\alpha u}(\tau,\xi) &= \left(1 + \frac{(\tau^2 - |\xi|^2)^2}{1 + \tau^2 + |\xi|^2}\right)^{\alpha/2} \widehat{u}(\tau,\xi).
\end{aligned}$$

It should be remarked that the weight of Λ_-^α is comparable to $\left(1 + \big||\tau| - |\xi|\big|\right)^\alpha$, but the former has the advantage of being smooth.

The Sobolev and "Wave Sobolev" spaces H^s and $H^{s,\theta}$ are given by the weighted L^2 norms

$$\|f\|_{H^s} = \|\Lambda^s f\|_{L^2(\mathbb{R}^4)} \quad \text{and} \quad \|u\|_{H^{s,\theta}} = \left\|\Lambda^s \Lambda_-^\theta u\right\|_{L^2(\mathbb{R}^{1+4})} .$$

We shall also use the related space $\mathcal{H}^{s,\theta}$ defined by

$$\|u\|_{\mathcal{H}^{s,\theta}} = \|u\|_{H^{s,\theta}} + \|\partial_t u\|_{H^{s-1,\theta}} \sim \left\|\Lambda^{s-1}\Lambda_+\Lambda_-^\theta u\right\|_{L^2}.$$

In view of Plancherel's theorem, these norms depend only on the size of the Fourier transform. It is an important fact that when $\theta > 1/2$, the spaces $H^{s,\theta}$ and $\mathcal{H}^{s,\theta}$ can be localized in time, since then the embeddings

$$H^{s,\theta} \hookrightarrow C_b(\mathbb{R}, H^s) \quad \text{and} \quad \mathcal{H}^{s,\theta} \hookrightarrow C_b(\mathbb{R}, H^s) \cap C_b^1(\mathbb{R}, H^{s-1}) \tag{17}$$

hold. See [7, Section 3].

Since $L^2(|\xi|^2\, d\xi) \subseteq L^1_{\text{loc}}(\mathbb{R}^4) \subseteq \mathcal{S}'(\mathbb{R}^4)$, we may define

$$\dot{H}^1 = \mathcal{F}^{-1}\left[L^2(|\xi|^2\, d\xi)\right].$$

Thus $\dot{H}^1$ is a Hilbert space with norm $\|f\|_{\dot{H}^1}^2 = \int_{\mathbb{R}^4} |\xi|^2 \left|\widehat{f}(\xi)\right|^2 d\xi$. We remark that if $\dot{W}^1 = \{f : \nabla f \in L^2\}$, then $\dot{H}^1$ is obtained by identifying elements of $\dot{W}^1$ differing by a constant. Observe also that $\mathcal{S}$ is dense in $L^2(|\xi|^2\, d\xi)$, hence in $\dot{H}^1$. We shall use frequently the fact that

$$\dot{H}^1 \hookrightarrow L^4(\mathbb{R}^4). \tag{18}$$

In other words, $\|f\|_{L^4} \lesssim \|f\|_{\dot{H}^1}$. This holds by the Hardy-Littlewood-Sobolev inequality (see Stein [14, Chapter V]).

If $\mathcal{X}$ is a separable Banach space of functions on $\mathbb{R}^4$, and $1 \le p \le \infty$, we denote by $L_t^p(\mathcal{X})$ the space $L^p(\mathbb{R}, \mathcal{X})$ of $\mathcal{X}$-valued functions. In particular, we write

$$\|u\|_{L_t^p(L_x^q)} = \left(\int_{\mathbb{R}} \|u(t,\cdot)\|_{L^q(\mathbb{R}^4)}^p\, dt\right)^{1/p}$$

with the usual modification if $p = \infty$.

We also need a version of this last norm which only depends on the size of the Fourier transform: If $u \in \mathcal{S}'$ and $\widehat{u}$ is a tempered function, set

$$\|u\|_{\mathcal{L}_t^p(\mathcal{L}_x^q)} = \sup\left\{\int_{\mathbb{R}^{1+4}} |\widehat{u}(\tau,\xi)|\, \widehat{v}(\tau,\xi)\, d\tau\, d\xi : v \in \mathcal{S},\, \widehat{v} \ge 0,\, \|v\|_{L_t^{p'}(L_x^{q'})} = 1\right\},$$

where $1 = \frac{1}{p} + \frac{1}{p'}$ and $1 = \frac{1}{q} + \frac{1}{q'}$. Let $\mathcal{L}_t^p(\mathcal{L}_x^q)$ be the corresponding subspace of $\mathcal{S}'$. Then $\|\cdot\|_{\mathcal{L}_t^p(\mathcal{L}_x^q)}$ is a translation invariant norm on $\mathcal{L}_t^p(\mathcal{L}_x^q)$. Note that $\mathcal{L}_t^2(\mathcal{L}_x^2) = L^2(\mathbb{R}^{1+4})$ and

$$\|u\|_{\mathcal{L}_t^p(\mathcal{L}_x^q)} \le \|u\|_{L_t^p(L_x^q)} \quad \text{whenever} \quad \widehat{u} \ge 0. \tag{19}$$

We refer the reader to [7, Section 4] for more details on these spaces.

We can now make precise the regularity statement (9). The solutions we obtain are in the following spaces:

$$A_0 \in C\big([0,T], \dot{H}^1\big) \cap C^1\big([0,T], L^2\big), \tag{20a}$$

$$A_j \in \mathcal{H}^{s,\theta} \cap \Lambda^{-\gamma}\Lambda_-^{-\frac{1}{2}}\left[\mathcal{L}_t^1(\mathcal{L}_x^8)\right], \tag{20b}$$

$$\phi \in \mathcal{H}^{s,\theta}, \tag{20c}$$

where $\theta > \frac{1}{2}$ and $\gamma > 0$ depend on s. For technical reasons, it is useful to iterate A_j and ϕ in these global spaces, but in the end we are only interested in their values on a time interval $[0, T]$ whose size depends on the norms of the data. Since the space $\mathcal{H}^{s,\theta}$ can be localized in time, this presents no problems.

Remark 1. The auxiliary space $\mathcal{L}_t^1(\mathcal{L}_x^8)$ in (20b) is necessary when $s < 5/4$. See Theorem 8.2 in [7] and the remark following it.

Note: Throughout the paper, we use the convenient shorthand $\lesssim$ for $\le$ up to a positive multiplicative constant C. Usually C is completely innocuous, and only depends on parameters that may be considered fixed. There are exceptions, notably for Lipschitz estimates (then C is only "locally" constant), but these are clearly pointed out.

1.5 Reformulation of the MKG system

As discussed in section 1.2, an important step in our proof is to recast the MKG system (6) as a system of nonlinear wave equations (11). Here we describe this in detail.

As was shown in [4], the first terms on the right hand sides of equations (6b) and (6c) can be expressed, due to the Coulomb condition (6d), in terms of the bilinear null forms

$$Q_{jk}(u, v) = \partial_j u \, \partial_k v - \partial_k u \, \partial_j v. \tag{21}$$

Since the argument in [4] was special to the case $n = 3$, we include here a proof of this fact which works for any dimension. First, let $\mathcal{P}$ be the projection onto the divergence free vector fields on $\mathbb{R}^4$. In terms of the Riesz transforms $R_j = (-\Delta)^{-\frac{1}{2}} \partial_j$,

$$\mathcal{P} X_j = X_j + R_j R^k X_k = R^k (R_j X_k - R_k X_j).$$

Observe that $\mathcal{P}$ is bounded on every L^p, $1 < p < \infty$, since this is true for the Riesz transforms (see Stein [14]). Moreover, it is clear that the Riesz transforms, and hence $\mathcal{P}$, are bounded on any space whose norm only depends on the size of the Fourier transform, in particular on any Sobolev space H^s.

Since $\partial_j(u \partial_k v) - \partial_k(u \partial_j v) = Q_{jk}(u, v)$, it follows immediately from the definition of $\mathcal{P}$ that

$$\mathcal{P}(u \partial_j v) = R^k (-\Delta)^{-\frac{1}{2}} Q_{jk}(u, v), \tag{22}$$

whence

$$\mathcal{P}\left(-\Im\left[\phi \overline{\partial_j \phi}\right]\right) = 2R^k (-\Delta)^{-\frac{1}{2}} Q_{jk}(\Re\phi, \Im\phi). \tag{23}$$

Also,

$$2\partial_j u \mathcal{P} X^j = Q_{jk}\left(u, (-\Delta)^{-\frac{1}{2}}\left[R^j X^k - R^k X^j\right]\right),$$

as one can see by expanding the right hand side. Therefore, if A is divergence free, so that $\mathcal{P}A = A$, then

$$2A^j\partial_j\phi = Q_{jk}\left(\phi, (-\Delta)^{-\frac{1}{2}}\left[R^jA^k - R^kA^j\right]\right). \tag{24}$$

Remark 2. The calculations leading to the identity (22) are certainly justified when u and v belong to the Schwartz class $\mathcal{S}(\mathbb{R}^4)$. Moreover, both sides of the identity are bounded bilinear operators of $(u,v) \in H^s \times H^s$ into H^{-1}, where $s > 1$. Thus the identity holds for all $u, v \in H^s$, and we conclude that (23) holds for all ϕ with the regularity (20c), since by (17) this implies $\phi \in C_b(\mathbb{R}, H^s)$. To bound the left hand side of (22), use first the dual

$$\left\|(-\Delta)^{-\frac{1}{2}}f\right\|_{L^2(\mathbb{R}^4)} \lesssim \|f\|_{L^{4/3}(\mathbb{R}^4)} \tag{25}$$

of (18). Since $\mathcal{P}$ is bounded on L^p, it then suffices to observe that

$$\|u\partial_jv\|_{L^{4/3}} \lesssim \|u\|_{L^4}\|\partial_ju\|_{L^2} \lesssim \|u\|_{H^1}\|v\|_{H^1}, \tag{26}$$

where we used (18). To prove boundedness of the right hand side of (22), it is enough to show

$$\left\|(-\Delta)^{-\frac{1}{2}}(fg)\right\|_{H^{-1}} \lesssim \|f\|_{H^{s-1}}\|g\|_{H^{s-1}}.$$

This can be reduced, via the self-duality of L^2, Plancherel's theorem, and the Cauchy-Schwarz inequality, to the fact that $|\xi|^{-1}(1+|\xi|)^{-1-2(s-1)}$ belongs to $L^2(\mathbb{R}^4)$, since $s > 1$. Similar, but simpler, considerations show that the remaining bilinear and cubic terms in (6a,b,c) and (28a,c,d) are bounded into $C_b\big(\mathbb{R}, L^{4/3}(\mathbb{R}^4)\big)$ when regarded as operators on A_0, A, ϕ in the class (20). For example, for a cubic expression uvw we have by Hölder's inequality and (18) that

$$\|uvw\|_{L^{4/3}} \le \|u\|_{L^4}\|v\|_{L^4}\|w\|_{L^4} \lesssim \|u\|_{\dot{H}^1}\|v\|_{\dot{H}^1}\|w\|_{\dot{H}^1}. \tag{27}$$

Returning to the main thread of our argument, we now use the null form identities derived above to arrive at an equivalent formulation of MKG:

$$\Delta A_0 = -\Im\big(\phi\overline{\partial_t\phi}\big) + |\phi|^2 A_0, \tag{28a}$$

$$\Delta\partial_t A_0 = -\Im\partial^j\big(\phi\overline{\partial_j\phi}\big) + \partial^j\big(|\phi|^2 A_j\big) \tag{28b}$$

$$\Box A_j = 2R^k(-\Delta)^{-\frac{1}{2}}Q_{jk}(\Re\phi, \Im\phi) + \mathcal{P}\big(|\phi|^2 A_j\big) \tag{28c}$$

$$\begin{aligned}\Box\phi = &-iQ_{jk}\left(\phi, (-\Delta)^{-\frac{1}{2}}\left[R^jA^k - R^kA^j\right]\right)\\ &+ 2iA_0\partial_t\phi + i(\partial_tA_0)\phi + A^\mu A_\mu\phi + m^2\phi.\end{aligned} \tag{28d}$$

This system acts as a stepping stone between (6) and (11).

Proposition 1. *The systems* (6) *and* (28) *are equivalent. More precisely, any local solution of* (6) *with the regularity* (20) *and divergence free initial data is a solution of* (28) *and vice versa.*

Proof. To go from (6) to (28), observe that A_j is divergence free by (6d); apply ∂^j to (6b) to get equation (28b); apply $\mathcal{P}$ to (6b) and use (23) to get (28c); finally, (28d) follows from (6c) using (24).

To go the other way, observe that by (23), the right hand side of (28c) is divergence free; thus $\Box \partial^j A_j = 0$, and since the initial data of A_j are divergence free, (6d) follows. Then, in view of (24), (6c) and (28d) are equivalent. Finally, to go from (28c) to (6b), it suffices to check that the right hand side of the latter is divergence free. But this follows from (28b). $\square$

Once the system has been written in the form (28) it is easy to eliminate A_0 and $\partial_t A_0$ and obtain the system of wave equations (11). We now describe this in more detail.

Lemma 1. *Given ϕ in the class* (9b)*, equation* (28a) *has a unique solution $A_0 \in \dot{H}^1$ on every time-slice $\{t\} \times \mathbb{R}^4$, and these solutions assemble to a space-time function $A_0 = A_0(\phi) \in C_b(\mathbb{R}, \dot{H}^1)$. Moreover, we have bounds, on every time-slice $\{t\} \times \mathbb{R}^4$,*

$$\|A_0\|_{\dot{H}^1} \leq 2 \|\partial_t \phi\|_{L^2}$$

and

$$\|A_0(\phi) - A_0(\psi)\|_{\dot{H}^1} \lesssim \|\phi - \psi\|_{H^1} + \|\partial_t \phi - \partial_t \psi\|_{L^2},$$

where the suppressed constant depends polynomially on $\|\phi\|_{H^1}, \|\psi\|_{H^1}, \|\partial_t \phi\|_{L^2}$ and $\|\partial_t \psi\|_{L^2}$, but is independent of t.

This is proved in section 4.

Next we consider (28b), with $\partial_t A_0$ replaced by the new variable B_0:

$$\Delta B_0 = -\Im \partial^j \big(\phi \overline{\partial_j \phi}\big) + \partial^j \big(|\phi|^2 A_j\big). \tag{29}$$

Lemma 2. *Given (A, ϕ) in the class* (9b)*, the equation* (29) *has a unique solution $B_0 \in L^2$ on every time-slice $\{t\} \times \mathbb{R}^4$, given by*

$$B_0 = R^j (-\Delta)^{-\frac{1}{2}} \Big[\Im \big(\phi \overline{\partial_j \phi}\big) - |\phi|^2 A_j \Big], \tag{30}$$

and the solutions assemble to a space-time function $B_0 = B_0(A, \phi) \in C_b(\mathbb{R}, L^2)$. Moreover, we have bounds, on every time-slice $\{t\} \times \mathbb{R}^4$,

$$\|B_0\|_{L^2} \leq C \big(1 + \|A\|_{H^1}\big) \|\phi\|_{H^1}^2$$

for a constant C independent of t, and

$$\|B_0(A, \phi) - B_0(A', \phi')\|_{L^2} \lesssim \|A - A'\|_{H^1} + \|\phi - \phi'\|_{H^1},$$

where the suppressed constant depends polynomially on $\|A\|_{H^1}, \|A'\|_{H^1}, \|\phi\|_{H^1}$ and $\|\phi'\|_{H^1}$, but is independent of t.

Proof. To see that (30) is in L^2, first apply (25), then estimate as in (26) and (27). That (30) is the only L^2 solution can be seen by taking the Fourier transform of both sides of (29). $\square$

In view of the above lemmas, (28) implies (11), with

$$\mathcal{M} = (\mathcal{M}_1, \dots, \mathcal{M}_4), \quad \mathcal{M}_j = \mathcal{M}_{j,1} + \mathcal{M}_{j,2}, \quad \mathcal{N} = \mathcal{N}_1 + \dots + \mathcal{N}_6,$$

where

$$\begin{aligned}
\mathcal{M}_{j,1} &= 2R^k(-\Delta)^{-\frac{1}{2}} Q_{jk}(\Re\phi, \Im\phi), \\
\mathcal{M}_{j,2} &= \mathcal{P}\big(|\phi|^2 A_j\big), \\
\mathcal{N}_1 &= -iQ_{jk}\left(\phi, (-\Delta)^{-\frac{1}{2}}\left[R^j A^k - R^k A^j\right]\right), \\
\mathcal{N}_2 &= 2iA_0(\phi)\partial_t\phi, \\
\mathcal{N}_3 &= iB_0(A,\phi)\phi, \\
\mathcal{N}_4 &= -[A_0(\phi)]^2\phi, \\
\mathcal{N}_5 &= |A|^2\phi, \\
\mathcal{N}_6 &= m^2\phi,
\end{aligned}$$

and $|A|^2 = A_j A^j$ in the next to last line.

Arguing as in Remark 2, and using Lemmas 1 and 2, it is readily checked that the multilinear expressions in $\mathcal{M}$ and $\mathcal{N}$ are all continuous maps into $C_b(\mathbb{R}, L^{4/3})$ [or $C_b(\mathbb{R}, H^{-1})$ in the case of $\mathcal{M}_{j,1}$] for (A,ϕ) in the class (20b,c). However, proving the following theorem requires much more sophisticated estimates.

Theorem 2. *The system of wave equations* (11)*, with $\mathcal{M}$ and $\mathcal{N}$ defined as above, is locally well-posed for initial data in H^s, all $s > 1$, in the following sense (all pairs (A,ϕ) are understood to belong to the class* (20b,c) *in what follows):*

- *(a)* ***(Local existence)*** *For all initial data* (7) *there exists a $T > 0$, which depends continuously on the norms of the data, and there exists a pair (A,ϕ) which solves* (11) *in the sense of distributions on $(0,T)\times\mathbb{R}^4$ and satisfies the given initial condition.*
- *(b)* ***(Uniqueness)*** *If $T > 0$ and we have two solutions (A,ϕ) and (A',ϕ') of* (11) *on $(0,T)\times\mathbb{R}^4$ with identical initial data, then they agree on the entire time-slab.*
- *(c)* ***(Continuous dependence on initial data)*** *If, for some $T > 0$, (A,ϕ) solves* (11) *on $(0,T)\times\mathbb{R}^4$ with initial data* (7)*, then for all initial data $(a', b', \phi'_0, \phi'_1)$ such that*

  $$\delta = \|a - a'\|_{H^s} + \|b - b'\|_{H^{s-1}} + \|\phi_0 - \phi'_0\|_{H^s} + \|\phi_1 - \phi'_1\|_{H^{s-1}}$$

  is sufficiently small, there is a solution (A',ϕ') on the same time-slab and with these initial data. Moreover, we have

  $$\|A - A'\|_{H^s} + \|\partial_t A - \partial_t A'\|_{H^{s-1}} + \|\phi - \phi'\|_{H^s} + \|\partial_t\phi - \partial_t\phi'\|_{H^{s-1}} \le C\delta$$

  uniformly in $0 \le t \le T$.

(d) ***(Persistence of higher regularity)*** *If k is a positive integer and (A,ϕ) solves* (11) *on $(0,T)\times\mathbb{R}^4$ with initial data in H^{s+k} (that is,* (7) *holds with s replaced by $s+k$), then*

$$A,\phi \in C\big([0,T],H^{s+k}\big)\cap C^1\big([0,T],H^{s+k-1}\big).$$

(e) ***(Classical solutions)*** *If the data belong to H^{s+k} for every k, then the solution is smooth:*

$$A,\phi \in C^\infty\big([0,T]\times\mathbb{R}^4\big).$$

The proof of this theorem will occupy us in the next two sections.

Here we want to show that Theorem 1 can be deduced from Theorem 2. It clearly suffices to demonstrate the equivalence of the systems (6) and (11). The remainder of this section is devoted to a proof of this fact, assuming that the conclusions of Theorem 2 hold.

Proposition 2. *The systems* (6) *and* (11) *are equivalent for local solutions in the regularity class* (20)*, with divergence free initial data.*

In view of Proposition 1, it suffices to show the equivalence of (28) and (11). We have seen already that (28) implies (11). The converse is not quite so obvious, but for sufficiently regular solutions it follows by some straightforward calulations and the fact, proved in section 4, that the only $\dot{H}^1$ solution of the elliptic equation $\Delta u = |\phi|^2 u$ is $u=0$. For general H^s data we then choose an approximating sequence of sufficiently regular data, use the persistence of higher regularity and continuous dependence on initial data, which hold by virtue of Theorem 2, and pass to the limit.

We now turn to the details.

Assume that (A,ϕ) is in the class (20b,c) and solves (11) on a time-slab $S_T=(0,T)\times\mathbb{R}^4$, with initial data satisfying (7) and (8). Set $A_0=A_0(\phi)$. Then (28) is satisfied, but with $\partial_t A_0$ replaced by $B_0=B_0(A,\phi)$ in (28b) and (28d). Thus, all we have to prove is that the distributional derivative $\partial_t A_0$ agrees with B_0 on S_T. At first glance one may think that this is simply a matter of taking a time derivative of (28a) and using the conservation law (4) to conclude that $\Delta\partial_t A_0 = \Delta B_0$, but this is a circular argument since the derivation of (4) is not valid unless we know that $\partial_t A_0 = B_0$.

In what follows, keep in mind that A_μ and B_0 are real-valued. Applying ∂_t to (28a) gives

$$\Delta\partial_t A_0 = -\Im\big(\phi\overline{\partial_t^2\phi}\big) + 2\Re\big(\phi\overline{\partial_t\phi}\big)A_0 + |\phi|^2\,\partial_t A_0. \tag{31}$$

Since (28c) and (8) hold, it follows as in the proof of Proposition 1 that A is divergence free. Therefore, (24) holds, and since (28d) holds (with $\partial_t A_0$ replaced by B_0), we conclude that

$$-\partial_t^2\phi + \Delta\phi = \Box\phi = -2iA^j\partial_j\phi + 2iA_0\partial_t\phi + iB_0\phi + A^\mu A_\mu\phi + m^2\phi.$$

Using this expression for $\partial_t^2 \phi$ gives, after some calculation,

$$-\Im\big(\phi\overline{\partial_t^2\phi}\big) = -\Im\partial^j\big(\phi\overline{\partial_j\phi}\big) + 2\Re\big(\phi\overline{\partial_j\phi}\big)A^j - 2\Re\big(\phi\overline{\partial_t\phi}\big)A_0 - B_0\left|\phi\right|^2.$$

Since

$$-\Im\partial^j\big(\phi\overline{\partial_j\phi}\big) = \Delta B_0 - \partial^j\big(\left|\phi\right|^2 A_j\big),$$

we get

$$-\Im\big(\phi\overline{\partial_t^2\phi}\big) = \Delta B_0 - \partial^j\big(\left|\phi\right|^2 A_j\big) + 2\Re\big(\phi\overline{\partial_j\phi}\big)A^j - 2\Re\big(\phi\overline{\partial_t\phi}\big)A_0 - B_0\left|\phi\right|^2.$$

Inserting this in (31) gives

$$\Delta\partial_t A_0 = \Delta B_0 - \partial^j\big(\left|\phi\right|^2 A_j\big) + 2\Re\big(\phi\overline{\partial_j\phi}\big)A^j - B_0\left|\phi\right|^2 + \left|\phi\right|^2\partial_t A_0.$$

But

$$\partial^j\big(\left|\phi\right|^2 A_j\big) = 2\Re\big(\phi\overline{\partial_j\phi}\big)A^j + \left|\phi\right|^2\partial^j A_j = 2\Re\big(\phi\overline{\partial_j\phi}\big)A^j$$

since A is divergence free, and so we finally get

$$\Delta(\partial_t A_0 - B_0) = \left|\phi\right|^2(\partial_t A_0 - B_0).$$

The above manipulations are justified provided

$$\partial_t A_0 \in C([0,T],\dot{H}^1). \tag{32}$$

If, moreover,

$$B_0 \in C([0,T],\dot{H}^1), \tag{33}$$

then it follows by the uniqueness result alluded to above (see Lemma 8 in section 4) that $\partial_t A_0 = B_0$ in $[0,T]\times\mathbb{R}^4$.

But (32) and (33) certainly hold under the additional assumption that the initial data (7) of A and ϕ belong to H^{s+k} for every positive integer k. Leaving aside the proof of this assertion for the moment, we note that any $f \in H^s$ can be approximated in the H^s norm by a sequence belonging to every H^{s+k}, by convolution with a C_c^∞ approximation of the identity, and if f is divergence free, then so is the approximating sequence. Combining these facts with the continuous dependence of A and ϕ on their H^s initial data (Theorem 2), and the continuity of the operators A_0 and B_0 (Lemmas 1 and 2), we conclude by passing to the limit that the equality $\partial_t A_0 = B_0$ holds in the sense of distributions on $(0,T)\times\mathbb{R}^4$ for all initial data (7) satisfying (8).

It remains to prove that (32) and (33) hold if the initial data (7) of A and ϕ belong to H^{s+k} for every positive integer k. For A_0, this follows by persistence of higher regularity (part (d) of Theorem 2), the inductive regularity step (73) in section 3.2 and Lemma 5 in the same section. As for B_0, in view of (30) it is clear that, on every time-slice,

$$\left\|B_0\right\|_{\dot{H}^1} \le \sum_j \Big(\left\|\phi\partial_j\phi\right\|_{L^2} + \left\|\left|\phi\right|^2 A_j\right\|_{L^2}\Big)$$

and by Hölder's inequality and Sobolev embedding it is easy to see that the right hand side is dominated by $\left\|\phi\right\|_{H^1}\left\|\phi\right\|_{H^2} + \left\|\phi\right\|_{H^2}^2\left\|A\right\|_{H^2}$. But if A and ϕ have initial data in H^{s+1}, then by persistence of higher regularity (part (d) of Theorem 2) we know that $A,\phi \in C([0,T],H^2)$.

2 Proof of Theorem 2

Here we discuss the estimates needed to prove local well-posedness of the system (11), with $\mathcal{M}$ and $\mathcal{N}$ defined as in section 1.5.

The local existence for the system (11) is proved by Picard iteration in the spaces (20b) and (20c), which are defined using the spacetime Fourier transform, and hence are global. However, since they embed in (9b), they can easily be localized in time. In fact, this time localization smooths out the singularity of the inverse $\square^{-1}$ of the wave operator, and — if done with sufficient care — allows one to handle large initial data by taking a sufficiently small time interval. These matters are considered in detail in the author's paper [12], and also in [7, Section 5], and we refer the interested reader there.

Fix $1 < s < 2$. (For larger s, the result can be proved by simpler arguments.) Let $\theta > \frac{1}{2}$ and $\gamma, \varepsilon > 0$; these quantities depend on the choice of s, and will be specified later. Now define

$$
\begin{aligned}
\mathcal{X}_1 &= \mathcal{H}^{s,\theta} \cap \Lambda^{-\gamma} \Lambda_-^{-\frac{1}{2}} \left[\mathcal{L}_t^1(\mathcal{L}_x^8) \right], \\
\mathcal{X}_2 &= \mathcal{H}^{s,\theta}, \\
\mathcal{Y}_k &= \Lambda_+ \Lambda_-^{1-\varepsilon} \mathcal{X}_k, \quad k = 1, 2
\end{aligned}
$$

with norms

$$
\begin{aligned}
\|A\|_{\mathcal{X}_1} &= \|A\|_{\mathcal{H}^{s,\theta}} + \left\| \Lambda^{\gamma} \Lambda_-^{\frac{1}{2}} A \right\|_{\mathcal{L}_t^1(\mathcal{L}_x^8)}, \\
\|\phi\|_{\mathcal{X}_2} &= \|\phi\|_{\mathcal{H}^{s,\theta}}, \\
\|F\|_{\mathcal{Y}_k} &= \left\| \Lambda_+^{-1} \Lambda_-^{-1+\varepsilon} F \right\|_{\mathcal{X}_k}, \quad k = 1, 2.
\end{aligned}
$$

All these spaces are complete (see [7, Proposition 4.2]), and by [7, Proposition 5.6], $\mathcal{X}_1$ and $\mathcal{X}_2$ satisfy the hypotheses of [12, Theorem 1]. Consequently, by [12, Theorem 2], the system (11) is locally well-posed for H^s data if the following Lipschitz conditions[9] hold:

$$
\|\mathcal{M}(A,\phi) - \mathcal{M}(A',\phi')\|_{\mathcal{Y}_1} \lesssim \|A - A'\|_{\mathcal{X}_1} + \|\phi - \phi'\|_{\mathcal{X}_2}, \tag{34a}
$$

$$
\|\mathcal{N}(A,\phi) - \mathcal{N}(A',\phi')\|_{\mathcal{Y}_2} \lesssim \|A - A'\|_{\mathcal{X}_1} + \|\phi - \phi'\|_{\mathcal{X}_2}, \tag{34b}
$$

where the suppressed constants depend continuously on

$$
\|A\|_{\mathcal{X}_1}, \quad \|A'\|_{\mathcal{X}_1}, \quad \|\phi\|_{\mathcal{X}_2} \quad \text{and} \quad \|\phi'\|_{\mathcal{X}_2}.
$$

In fact, these estimates guarantee that the conclusions (a,b,c) of Theorem 2 hold. In the next section we show how to prove parts (d) and (e) of the same theorem.

It suffices to prove (34) with $\mathcal{M}$ replaced by $\mathcal{M}_{j,k}$ and with $\mathcal{N}$ replaced by $\mathcal{N}_1, \dots, \mathcal{N}_5$. Furthermore, in view of the multilinear structure, it suffices to

[9] Keep in mind that $\mathcal{M}$ and $\mathcal{N}$ vanish at the origin, so if we take $A' = 0$ and $\phi' = 0$, we simply get bounds for $\mathcal{M}(A,\phi)$ and $\mathcal{N}(A,\phi)$.

prove (concerning the suppressed constants, see note below):

$$\|\mathcal{M}_{j,1}\|_{\mathcal{Y}_1} \lesssim \|\phi\|_{\mathcal{X}_2}^2 , \tag{35}$$
$$\|\mathcal{M}_{j,2}\|_{\mathcal{Y}_1} \lesssim \|A\|_{\mathcal{X}_1} \|\phi\|_{\mathcal{X}_2}^2 , \tag{36}$$
$$\|\mathcal{N}_1\|_{\mathcal{Y}_2} \lesssim \|A\|_{\mathcal{X}_1} \|\phi\|_{\mathcal{X}_2} , \tag{37}$$
$$\|\mathcal{N}_2\|_{\mathcal{Y}_2} \lesssim \|A_0(\phi)\|_{\mathcal{Z}_1} \|\phi\|_{\mathcal{X}_2} , \tag{38}$$
$$\|\mathcal{N}_3\|_{\mathcal{Y}_2} \lesssim \|B_0(A,\phi)\|_{\mathcal{Z}_2} \|\phi\|_{\mathcal{X}_2} , \tag{39}$$
$$\|\mathcal{N}_4\|_{\mathcal{Y}_2} \lesssim \|A_0(\phi)\|_{L_t^\infty(\dot{H}^1)} \|A_0(\phi)\|_{\mathcal{Z}_1} \|\phi\|_{\mathcal{X}_2} , \tag{40}$$
$$\|\mathcal{N}_5\|_{\mathcal{Y}_2} \lesssim \|A\|_{\mathcal{X}_1}^2 \|\phi\|_{\mathcal{X}_2} , \tag{41}$$
$$\|\mathcal{N}_6\|_{\mathcal{Y}_2} \le \|\phi\|_{\mathcal{X}_2} , \tag{42}$$
$$\|A_0\|_{L_t^\infty(\dot{H}^1)} \lesssim \|\phi\|_{\mathcal{X}_2} , \tag{43}$$
$$\|A_0(\phi) - A_0(\phi')\|_{L_t^\infty(\dot{H}^1)} \lesssim \|\phi - \phi'\|_{\mathcal{X}_2} , \tag{44}$$

where $\mathcal{Z}_1$ and $\mathcal{Z}_2$ are certain intermediate spaces, to be specified later, such that

$$\|A_0\|_{\mathcal{Z}_1} \lesssim \|\phi\|_{\mathcal{X}_2}^2 + \|\phi\|_{\mathcal{X}_2}^3 , \tag{45}$$
$$\|A_0(\phi) - A_0(\phi')\|_{\mathcal{Z}_1} \lesssim \|\phi - \phi'\|_{\mathcal{X}_2} , \tag{46}$$
$$\|B_0\|_{\mathcal{Z}_2} \lesssim \left(1 + \|A\|_{\mathcal{X}_1}\right) \|\phi\|_{\mathcal{X}_2}^2 , \tag{47}$$
$$\|B_0(A,\phi) - B_0(A',\phi')\|_{\mathcal{Z}_2} \lesssim \|A - A'\|_{\mathcal{X}_1} + \|\phi - \phi'\|_{\mathcal{X}_2} . \tag{48}$$

It should be emphasized that in the Lipschitz estimates (44), (46) and (48), the suppressed constant depends polynomially on the norms $\|\phi\|_{\mathcal{X}_2}$ and $\|\phi'\|_{\mathcal{X}_2}$, and in the case of (48) also on $\|A\|_{\mathcal{X}_1}$ and $\|A'\|_{\mathcal{X}_1}$. Observe that the estimate (42) for the linear term is trivial, since the norms only depend on the size of the Fourier transform.

The following was proved in [7, Theorem 8.6].

Theorem. *The estimates* (35) *and* (37) *hold provided*

$$\frac{1}{2} < \theta < \min\left(\frac{3}{4}, \frac{s}{2}\right) \tag{49a}$$
$$0 < \varepsilon < \frac{1}{4} \min\left(\frac{3}{4} - \theta, \frac{s}{2} - \theta\right) \tag{49b}$$
$$\gamma = \theta - \frac{1}{2} - 3\varepsilon. \tag{49c}$$

Having fixed θ and ε satisfying these requirements, we define p and r by

$$\frac{1}{p} = \frac{3}{2} - \theta - 2\varepsilon, \quad \frac{1}{r} = 1 - \theta - 2\varepsilon, \tag{50}$$

and we choose q so large that

$$\frac{4}{q} < \min\left(2\theta - 1, 1 - \frac{1}{p}\right). \tag{51}$$

Observe that as $s \to 1$, the triple $(p,q,r) \to (1,\infty,2)$. Now set

$$\|A_0\|_{\mathcal{Z}_1} = \left\|\Lambda^{s-1} A_0\right\|_{L_t^p(L_x^q)}, \tag{52}$$

$$\|B_0\|_{\mathcal{Z}_2} = \left\|\Lambda^{s-1} B_0\right\|_{L_t^r(L_x^{8/3})}. \tag{53}$$

For easy reference, we list here some estimates that we shall use (here p,q,r are defined as above):

$$\left\|\Lambda^{s-1}(-\Delta)^{-1}(uv)\right\|_{L_t^p(L_x^q)} \lesssim \|u\|_{H^{s,\theta}} \|v\|_{H^{s-1,\theta}}, \tag{54}$$

$$\|u\|_{L_t^2(L_x^8)} \lesssim \|u\|_{H^{1,\theta}}, \tag{55}$$

$$\|u\|_{L_t^r(L_x^8)} \lesssim \|u\|_{H^{s,\theta}}, \tag{56}$$

$$\|u\|_{L_t^{2p}(L_x^\beta)} \lesssim \|u\|_{H^{s,\theta}}, \quad \frac{5}{2} + \theta + 2\varepsilon - 2s \le \frac{8}{\beta} \le 2\theta, \tag{57}$$

$$\|u\|_{H^{0,\theta+\varepsilon-1}} \lesssim \|u\|_{L_t^p(L_x^2)}, \tag{58}$$

$$\|u\|_{\mathcal{L}_t^1(\mathcal{L}_x^8)} \lesssim \left\|\Lambda\Lambda_-^{\frac{1}{2}+\varepsilon} u\right\|_{\mathcal{L}_t^1(\mathcal{L}_x^2)}, \tag{59}$$

$$\|fg\|_{H^\sigma} \lesssim \|\Lambda^\sigma f\|_{L^{p_1}} \|g\|_{L^{q_1}} + \|\Lambda^\sigma g\|_{L^{p_2}} \|f\|_{L^{q_2}}, \tag{60}$$

where in the last inequality,

$$\sigma > 0, \quad \frac{1}{p_k} + \frac{1}{q_k} = \frac{1}{2}, \quad 2 \le p_k < \infty.$$

The inequality (54) follows from a theorem of Klainerman-Tataru [8]; we give the details in an appendix.

The Strichartz type estimates (55–57) are special cases of [7, Theorem D]. (The [non-optimal] upper bound for $8/\beta$ in (57) guarantees that the pair $(2p,\beta)$ is wave admissible; the lower bound is chosen so that we do not exceed s space derivatives on the right hand side.)

The inequality (58) can either be proved directly, using Plancherel's theorem, Hölder's inequality, Minkowski's integral inequality and the Hausdorff-Young inequality, or it can be proved by interpolation, as in [7, Section 6(vii)]).

Inequality (59) is a special case of [7, Proposition 4.8].

The calculus inequality (60) is Lemma 1 in Ponce-Sideris [9].

As mentioned already, (35) and (37) hold by [7, Theorem 8.6]. We now prove the remaining estimates (36) and (38–48), thereby concluding the proof of parts (a,b,c) of Theorem 2.

2.1 Proof of (36)

Since the norm only depends on the size of the Fourier transform, we can ignore the projection $\mathcal{P}$. More accurately,

$$\|\mathcal{M}_{j,2}\|_{\mathcal{Y}_1} \lesssim \left\||\phi|^2 A_j\right\|_{\mathcal{Y}_1}.$$

Thus, it suffices to prove

$$\left\|\Lambda_+^{-1}\Lambda_-^{\varepsilon-1}(uvw)\right\|_{\mathcal{X}_1} \lesssim \|u\|_{\mathcal{H}^{s,\theta}} \|v\|_{\mathcal{H}^{s,\theta}} \|w\|_{\mathcal{H}^{s,\theta}} ,$$

or, equivalently,

$$\|uvw\|_{H^{s-1,\theta+\varepsilon-1}} \lesssim \|u\|_{\mathcal{H}^{s,\theta}} \|v\|_{\mathcal{H}^{s,\theta}} \|w\|_{\mathcal{H}^{s,\theta}} ,$$
$$\left\|\Lambda^\gamma\Lambda_+^{-1}\Lambda_-^{\varepsilon-\frac{1}{2}}(uvw)\right\|_{\mathcal{L}_t^1(\mathcal{L}_x^8)} \lesssim \|u\|_{\mathcal{H}^{s,\theta}} \|v\|_{\mathcal{H}^{s,\theta}} \|w\|_{\mathcal{H}^{s,\theta}} .$$

Since all the norms depend only on the size of the Fourier transform, we may assume that u, v, w have non-negative Fourier transforms, and we see that it is sufficient to prove (note that $\gamma + 2\varepsilon < s - 1$ by (49))

$$\|uvw\|_{H^{0,\theta+\varepsilon-1}} \lesssim \|u\|_{\mathcal{H}^{1,\theta}} \|v\|_{\mathcal{H}^{s,\theta}} \|w\|_{\mathcal{H}^{s,\theta}} , \tag{61}$$
$$\left\|\Lambda^{-1}\Lambda_-^{-\varepsilon-\frac{1}{2}}(uvw)\right\|_{\mathcal{L}_t^1(\mathcal{L}_x^8)} \lesssim \|u\|_{\mathcal{H}^{1,\theta}} \|v\|_{\mathcal{H}^{s,\theta}} \|w\|_{\mathcal{H}^{s,\theta}} . \tag{62}$$

By (58) and Hölder's inequality,

$$\|uvw\|_{H^{0,\theta+\varepsilon-1}} \lesssim \|u\|_{L_t^\infty(L_x^4)} \|v\|_{L_t^{2p}(L_x^8)} \|w\|_{L_t^{2p}(L_x^8)} ,$$

and (61) follows by Sobolev embedding and (57).

Using (59) and (19), we get

$$\left\|\Lambda^{-1}\Lambda_-^{-\varepsilon-\frac{1}{2}}(uvw)\right\|_{\mathcal{L}_t^1(\mathcal{L}_x^8)} \lesssim \|uvw\|_{L_t^1(L_x^2)} \lesssim \|u\|_{L_t^\infty(L_x^4)} \|v\|_{L_t^2(L_x^8)} \|w\|_{L_t^2(L_x^8)} .$$

Now use Sobolev embedding and (55).

2.2 Proof of (38)

We have to show

$$\|uv\|_{H^{s-1,\theta+\varepsilon-1}} \lesssim \left\|\Lambda^{s-1}u\right\|_{L_t^p(L_x^q)} \|v\|_{\mathcal{H}^{s-1,\theta}} .$$

By (58) and (60),

$$\begin{aligned}\|uv\|_{H^{s-1,\theta+\varepsilon-1}} &\lesssim \left\|\Lambda^{s-1}(uv)\right\|_{L_t^p(L_x^2)} \\ &\lesssim \left\|\Lambda^{s-1}u\right\|_{L_t^p(L_x^q)} \|v\|_{L_t^\infty(L_x^{(1/2-1/q)^{-1}})} + \|u\|_{L_t^p(L_x^\infty)} \left\|\Lambda^{s-1}v\right\|_{L_t^\infty(L_x^2)} .\end{aligned}$$

The desired estimate now follows by Sobolev embedding, since $\frac{4}{q} < s - 1$.

2.3 Proof of (39)

We must prove

$$\|uv\|_{H^{s-1,\theta+\varepsilon-1}} \lesssim \left\|\Lambda^{s-1}u\right\|_{L_t^r(L_x^{8/3})} \|v\|_{\mathcal{H}^{s,\theta}} .$$

By (58) and (60),

$$\begin{aligned}\|uv\|_{H^{s-1,\theta+\varepsilon-1}} &\lesssim \left\|\Lambda^{s-1}(uv)\right\|_{L_t^p(L_x^2)} \\ &\lesssim \left\|\Lambda^{s-1}u\right\|_{L_t^2(L_x^{8/3})} \|v\|_{L_t^r(L_x^8)} + \|u\|_{L_t^r(L_x^{8/3})} \left\|\Lambda^{s-1}v\right\|_{L_t^2(L_x^8)} . \qquad (63)\end{aligned}$$

Now apply (55) and (56). Note also that $\|u\|_{L_t^r(L_x^{8/3})} \lesssim \left\|\Lambda^{s-1}u\right\|_{L_t^r(L_x^{8/3})}$, since $\Lambda^{-\delta}$ is bounded on L^p for all $1 \le p \le \infty$ and $\delta \ge 0$. In fact, $\Lambda^{-\delta}$ corresponds to convolution with an L^1 function; see Stein [14].

2.4 Proof of (40)

It suffices to show

$$\left\|u^2v\right\|_{H^{s-1,\theta+\varepsilon-1}} \lesssim \|u\|_{L_t^\infty(L_x^4)} \left\|\Lambda^{s-1}u\right\|_{L_t^p(L_x^q)} \|v\|_{\mathcal{H}^{s,\theta}} .$$

By (58) and (60),

$$\begin{aligned}\left\|u^2v\right\|_{H^{s-1,\theta+\varepsilon-1}} &\lesssim \left\|\Lambda^{s-1}(u^2v)\right\|_{L_t^p(L_x^2)} \\ &\lesssim \left\|\Lambda^{s-1}u\right\|_{L_t^p(L_x^q)} \|uv\|_{L_t^\infty(L_x^{(1/2-1/q)^{-1}})} + \|u\|_{L_t^\infty(L_x^4)} \left\|\Lambda^{s-1}(uv)\right\|_{L_t^p(L_x^4)} \\ &\lesssim \left\|\Lambda^{s-1}u\right\|_{L_t^p(L_x^q)} \|u\|_{L_t^\infty(L_x^4)} \|v\|_{L_t^\infty(L_x^{(1/4-1/q)^{-1}})} \\ &\qquad + \|u\|_{L_t^\infty(L_x^4)} \|u\|_{L_t^p(L_x^\infty)} \left\|\Lambda^{s-1}v\right\|_{L_t^\infty(L_x^4)} .\end{aligned}$$

Now apply Sobolev embedding, and use (51).

2.5 Proof of (41)

It suffices to show

$$\|uvw\|_{H^{s-1,\theta+\varepsilon-1}} \lesssim \|u\|_{\mathcal{H}^{s,\theta}} \|v\|_{\mathcal{H}^{s,\theta}} \|w\|_{\mathcal{H}^{s,\theta}} ,$$

but this was proved above; see the proof of (36).

2.6 Proof of (43) and (44)

These follow from Lemma 1, which is proved in section 4.

2.7 Proof of (45) and (46)

Since

$$A_0 = (-\Delta)^{-1} \left[\Im\big(\phi\overline{\partial_t\phi}\big) - |\phi|^2 A_0 \right],$$

it suffices, taking into account the multilinearity of the terms inside the brackets, as well as the estimates (43) and (44), to show that

$$\begin{aligned}\left\|\Lambda^{s-1}(-\Delta)^{-1}(uv)\right\|_{L_t^p(L_x^q)} &\lesssim \|u\|_{H^{s,\theta}} \|v\|_{H^{s-1,\theta}} , \\ \left\|\Lambda^{s-1}(-\Delta)^{-1}(uvw)\right\|_{L_t^p(L_x^q)} &\lesssim \|u\|_{H^{s,\theta}} \|v\|_{H^{s,\theta}} \|w\|_{L_t^\infty(L_x^4)} .\end{aligned}$$

The former is exactly (54), and the left hand side of the latter is $\lesssim$

$$\left\|(-\Delta)^{-1}(uvw)\right\|_{L_t^p(L_x^q)} + \left\|(-\Delta)^{\frac{s-3}{2}}(uvw)\right\|_{L_t^p(L_x^q)}. \tag{64}$$

Here we applied the following useful result, which is an immediate consequence of Lemma 2(ii) in Chapter V of Stein [14].

Lemma 3. *For $\alpha > 0$ and $1 \le p \le \infty$,*

$$\|\Lambda^\alpha f\|_{L^p} \lesssim \|f\|_{L^p} + \left\|(-\Delta)^{\alpha/2} f\right\|_{L^p},$$

where the suppressed constant only depends on α.

Returning to the sum (64), note that by Sobolev embedding, it is $\lesssim$

$$\|uvw\|_{L_t^p(L_x^{\alpha_1})} + \left\|uvw\right\|_{L_t^p(L_x^{\alpha_2})},$$

where

$$\frac{1}{\alpha_1} = \frac{1}{2} + \frac{1}{q} = \frac{1}{4} + 2\left(\frac{1}{8} + \frac{1}{2q}\right),$$
$$\frac{1}{\alpha_2} = \frac{3-s}{4} + \frac{1}{q} = \frac{1}{4} + 2\left(\frac{2-s}{8} + \frac{1}{2q}\right).$$

Thus

$$\|uvw\|_{L_t^p(L_x^{\alpha_k})} \le \|u\|_{L_t^{2p}(L_x^{\beta_k})} \|v\|_{L_t^{2p}(L_x^{\beta_k})} \|w\|_{L_t^\infty(L_x^4)}, \quad k = 1, 2$$

where

$$\frac{1}{\beta_1} = \frac{1}{8} + \frac{1}{2q}, \quad \frac{1}{\beta_2} = \frac{2-s}{8} + \frac{1}{2q}.$$

Using (49b) and (51) it is easily checked that

$$\frac{5}{2} + \theta + 2\varepsilon - 2s \le 2 - s \le \frac{8}{\beta_2} < \frac{8}{\beta_1} < 2\theta,$$

so we may apply (57) to finish the proof.

2.8 Proof of (47) and (48)

We prove (47); the same proof gives (48) if one exploits the multilinearity of the terms defining B_0.

First observe that by Lemma 3,

$$\left\|\Lambda^{s-1} B_0\right\|_{L_t^r(L_x^{8/3})} \lesssim \|B_0\|_{L_t^r(L_x^{8/3})} + \left\|(-\Delta)^{\frac{s-1}{2}} B_0\right\|_{L_t^r(L_x^{8/3})}.$$

Therefore, by Sobolev embedding, we have to estimate

$$\left\|(-\Delta)^{\frac{1}{2}} B_0\right\|_{L_t^r(L_x^{\alpha_k})}, \quad k = 1, 2$$

where

$$\frac{1}{\alpha_1} = \frac{5}{8}, \quad \frac{1}{\alpha_2} = \frac{5}{8} - \frac{s-1}{4}.$$

Since B_0 is given by (30), and since the Riesz transforms R_j are bounded on L^p, $1 < p < \infty$, we see that it is enough to prove

$$\begin{aligned} \|uv\|_{L_t^r(L_x^{\alpha_k})} &\lesssim \|u\|_{H^{s,\theta}} \|v\|_{H^{s-1,\theta}}, \\ \|uvw\|_{L_t^r(L_x^{\alpha_k})} &\lesssim \|u\|_{H^{s,\theta}} \|v\|_{H^{s,\theta}} \|w\|_{H^{s,\theta}}. \end{aligned}$$

By Hölder's inequality,

$$\begin{aligned} \|uv\|_{L_t^r(L_x^{\alpha_k})} &\le \|u\|_{L_t^r(L_x^8)} \|v\|_{L_t^\infty(L_x^{\beta_k})}, \\ \|uvw\|_{L_t^r(L_x^{\alpha_k})} &\le \|u\|_{L_t^r(L_x^8)} \|v\|_{L_t^\infty(L_x^4)} \|v\|_{L_t^\infty(L_x^{\gamma_k})}, \end{aligned}$$

where

$$\begin{aligned} &\beta_1 = 2, & &\gamma_1 = 4, \\ &\frac{1}{\beta_2} = \frac{1}{2} - \frac{s-1}{4}, & &\frac{1}{\gamma_2} = \frac{1}{4} - \frac{s-1}{4}. \end{aligned}$$

Now apply (56) and Sobolev embedding.

3 Higher regularity

Here we prove parts (d) and (e) of Theorem 2.

3.1 The persistence property

The key to proving part (d) of Theorem 2 is to establish, for $k = 0, 1, 2, \dots$,

$$\left\|\Lambda^k \mathcal{M}(A,\phi)\right\|_{\mathcal{Y}_1} \le \alpha_k \left\{ \left\|\Lambda^k A\right\|_{\mathcal{X}_1} + \left\|\Lambda^k \phi\right\|_{\mathcal{X}_2} \right\} + \beta_k, \tag{65a}$$

$$\left\|\Lambda^k \mathcal{N}(A,\phi)\right\|_{\mathcal{Y}_2} \le \alpha_k \left\{ \left\|\Lambda^k A\right\|_{\mathcal{X}_1} + \left\|\Lambda^k \phi\right\|_{\mathcal{X}_2} \right\} + \beta_k, \tag{65b}$$

where

- α_k depends continuously on $\|A\|_{\mathcal{X}_1}$ and $\|\phi\|_{\mathcal{X}_2}$,
- $\beta_0 = 0$,
- β_k, for $k \ge 1$, depends continuously on $\left\|\Lambda^{k-1} A\right\|_{\mathcal{X}_1}$ and $\left\|\Lambda^{k-1}\phi\right\|_{\mathcal{X}_2}$.

The case $k = 0$ is of course true by (34), but it is useful to include it here for technical reasons.

In the absence of the lower order term β_k, we could now appeal directly to [12, Theorem 2], to conclude that part (d) of Theorem 2 holds. However, we

can easily modify the proof given in [12] to cover this more general case, as we demonstrate below.

First, however, let us dispose of proof of the above estimates. Observe that we have the equivalence of norms

$$\left\|\Lambda^k u\right\|_{\mathcal{X}_j} \sim \sum_{|\alpha|\le k} \|\partial_x^\alpha u\|_{\mathcal{X}_j} .$$

This is trivial in view of the fact that the norms only depend on the size of the Fourier transform. It is therefore clear, from the multilinear structure of $\mathcal{M}$ and $\mathcal{N}$, and the product rule for derivatives, that (65) follows from the very estimates proved in section 2. The only exception is the nonlinear operator $A_0(\phi)$, for which we need the following estimate, replacing (43):

Lemma 4. *If $\Lambda^k\phi \in \mathcal{X}_2$, then*

$$\|\partial_x^\alpha A_0\|_{L_t^\infty(\dot H^1)} \le \gamma_k\left(\|\phi\|_{\mathcal{X}_2}\right)\left\|\Lambda^k\phi\right\|_{\mathcal{X}_2} + \eta_k\left(\left\|\Lambda^{k-1}\phi\right\|_{\mathcal{X}_2}\right) \quad \textit{for all} \quad |\alpha| \le k,$$

where γ_k and η_k are continuous functions.

This is proved in section 4.3.

Let us now turn to the proof of Theorem 2, part (d).

The issue is to show that if we have a pair (A, ϕ), belonging to the class (20b,c), which solves (11) on $S_T = (0, T) \times \mathbb{R}^4$ with initial data (7), and if the data have some additional regularity, say H^{s+k}, then this extra regularity persists throughout the time interval $[0, T]$:

$$A, \phi \in C\big([0,T], H^{s+k}\big) \cap C^1\big([0,T], H^{s+k-1}\big). \tag{66}$$

Now, as proved in [12, Section 6.4], it suffices to prove this for *some* $T > 0$ *which depends continuously on*

$$E_0 = \|a\|_{H^s} + \|b\|_{H^{s-1}} + \|\phi_0\|_{H^s} + \|\phi_1\|_{H^{s-1}} .$$

We shall prove this using the Picard iterates corresponding to the given data. It will be convenient to introduce the notation

$$E_k = \|a\|_{H^{s+k}} + \|b\|_{H^{s+k-1}} + \|\phi_0\|_{H^{s+k}} + \|\phi_1\|_{H^{s+k-1}} .$$

Now fix an integer $K \ge 1$, and denote by α and β the pointwise maxima of α_k and β_k, respectively, taken over all $0 \le k \le K$. Let us assume that the initial data belong to H^{s+K}, that is,

$$E_K < \infty.$$

It is proved in [12] that for any $0 < T < 1$, there is a linear operator W_T, which is bounded from $\mathcal{Y}_j \to \mathcal{X}_j$ ($j = 1, 2$), and such that $u = W_T F$ solves the inhomogeneous wave equation $\Box u = F$ on $(0, T) \times \mathbb{R}^4$ with vanishing initial data at $t = 0$. Moreover, if C_T is the maximum of the operator norms, that is,

$$C_T = \max\left(\|W_T\|_{\mathcal{Y}_1\to\mathcal{X}_1}, \|W_T\|_{\mathcal{Y}_2\to\mathcal{X}_2}\right), \tag{67}$$

then

$$C_T \to 0 \quad \text{as} \quad T \to 0. \tag{68}$$

The sequence of Picard iterates $(A^{(m)}, \phi^{(m)})$ is then defined inductively as follows. First, let $A^{(0)}$ and $\phi^{(0)}$ be the solutions of $\Box A^{(0)} = 0$ and $\Box \phi^{(0)} = 0$ with initial data (7), and then multiply them by a smooth bump function which equals 1 on the interval $[0, T]$. By [12, Theorem 1],

$$\left\| \Lambda^k A^{(0)} \right\|_{\mathcal{X}_1} + \left\| \Lambda^k \phi^{(0)} \right\|_{\mathcal{X}_2} \le C E_k, \tag{69}$$

with E_k as above. Then define

$$\begin{aligned} A^{(m+1)} &= A^{(0)} + W_T \mathcal{M}(A^{(m)}, \phi^{(m)}), \\ \phi^{(m+1)} &= \phi^{(0)} + W_T \mathcal{N}(A^{(m)}, \phi^{(m)}). \end{aligned}$$

Let us write

$$\begin{aligned} R_k^{(m)} &= \left\| \Lambda^k A^{(m)} \right\|_{\mathcal{X}_1} + \left\| \Lambda^k \phi^{(m)} \right\|_{\mathcal{X}_2}, \\ \omega^{(m)} &= \left\| A^{(m)} - A^{(m-1)} \right\|_{\mathcal{X}_1} + \left\| \phi^{(m)} - \phi^{(m-1)} \right\|_{\mathcal{X}_2}. \end{aligned}$$

Then by (69), (67) and (65) (with $k = 0$), we have

$$R_0^{(m+1)} \le C E_0 + C_T \, \alpha\big(R_0^{(m)}\big) R_0^{(m)}, \quad m \ge 0.$$

If we choose T so small that

$$2 C_T \, \alpha(2 C E_0) \le 1, \tag{70}$$

then it follows by induction on m that

$$R_0^{(m)} \le 2 C E_0, \quad m \ge 0. \tag{71}$$

Then, using the Lipschitz estimates (34) (and making α larger if necessary),

$$\omega^{(m+1)} \le \frac{1}{2} \omega^{(m)},$$

so the sequence of Picard iterates is Cauchy in $\mathcal{X}_1 \times \mathcal{X}_2$, and therefore converges; the limit is of course the unique solution (A, ϕ) of our equation.

We shall prove that, *with T as in* (70)*,*

$$R_k^{(m)} \le C_k(E_0, \ldots, E_k), \quad k \le K, \quad m \ge 0, \tag{72}$$

where C_k is some continuous function.

Let us first see why this implies the desired conclusion (66) for $k \le K$. The point is that by (72), the sequence of Picard iterates is bounded in the Hilbert space $\mathcal{H}^{s+k,\theta}$ (recall that $\mathcal{X}_1 \hookrightarrow \mathcal{X}_2 = \mathcal{H}^{s,\theta}$), and therefore, some subsequence

converges weakly in that space. Since weak convergence in $\mathcal{H}^{s+k,\theta}$ implies convergence in the sense of distributions, we conclude that the strong limit (A, ϕ) agrees, as a distribution, with this weak limit. Thus, (A, ϕ) belongs to $\mathcal{H}^{s+k,\theta}$, and this immediately gives (66).

We shall prove (72) by induction on k.

We already have the case $k = 0$, by (71).

Now assume that $k < K$ and that (72) holds. We claim that this implies (72) for $k + 1$. Indeed, by (65), (67) and (69),

$$R_{k+1}^{(m+1)} \le CE_{k+1} + C_T \alpha\big(R_0^{(m)}\big) R_{k+1}^{(m)} + C_T \beta\big(R_k^{(m)}\big).$$

Taking into account (71), (70) and the induction hypothesis, we get

$$R_{k+1}^{(m+1)} \le CE_{k+1} + \frac{1}{2} R_{k+1}^{(m)} + \frac{\beta\big(C_k(E_0, \dots, E_k)\big)}{2\alpha(2CE_0)}$$

for $m \ge 0$. It now follows by induction on m that

$$R_{k+1}^{(m)} \le 2CE_{k+1} + \frac{\beta\big(C_k(E_0, \dots, E_k)\big)}{\alpha(2CE_0)}, \quad m \ge 0,$$

using (69) for the case $m = 0$.

3.2 Classical solutions

Here we outline the proof of part (e) of Theorem 2. In view of part (d) of the same theorem, it suffices to prove the inductive step

$$A, \phi \in \bigcap_{k=1}^{\infty} C^m\big([0, T], H^{s+k}\big) \implies A, \phi \in \bigcap_{k=1}^{\infty} C^{m+1}\big([0, T], H^{s+k}\big). \tag{73}$$

But since (A, ϕ) solves (11) on $(0, T) \times \mathbb{R}^4$, we have there

$$\partial_t^2 A = \Delta A - \mathcal{M}(A, \phi),$$
$$\partial_t^2 \phi = \Delta \phi - \mathcal{N}(A, \phi),$$

and so it is clear that (73) follows from

$$A, \phi \in \bigcap_{k=1}^{\infty} C^m\big([0, T], H^{s+k}\big) \\ \implies \mathcal{M}(A, \phi), \mathcal{N}(A, \phi) \in \bigcap_{k=1}^{\infty} C^{m-1}\big([0, T], H^{s+k}\big). \tag{74}$$

The key observation is of course that $\mathcal{M}$ and $\mathcal{N}$ only contain first order derivatives in time. Recall that $\mathcal{M}$ and $\mathcal{N}$ are sums of multilinear expressions in A and ϕ and their first order derivatives, and terms involving $A_0(\phi)$. But

$A_0(\phi)$ is determined by the elliptic equation (28a), which also contains only first order partial derivatives in time of ϕ.

Thus, to prove (74), simply apply up to $m-1$ time derivatives and any number of space derivatives, say K, to $\mathcal{M}$ and $\mathcal{N}$, and use the product rule for derivatives. It is then easy to show — we omit the details — that on each time-slice, the L^2-norms of the resulting expressions are bounded in terms of (here α is a multi-index)

$$\left\|\partial_t^j A\right\|_{H^{K+k}}, \quad \left\|\partial_t^j \phi\right\|_{H^{K+k}} \quad \text{and} \quad \left\|\partial_t^j \partial_x^\alpha A_0(\phi)\right\|_{\dot{H}^1}$$

for $j \le m$, $|\alpha| \le K$ and k sufficiently large. Then one appeals to the following higher regularity result for $A_0(\phi)$, which is proved in section 4.3.

Lemma 5. *Let m, M be non-negative integers. If $\phi \in C^{m+1}\big([0,T], H^M\big)$, that is, if*

$$\partial_t^j \partial_x^\alpha \phi \in C\big([0,T], L^2\big) \quad \textit{for all} \quad j \le m+1 \quad \textit{and all} \quad |\alpha| \le M+1,$$

where α is a multi-index, then

$$\partial_t^j \partial_x^\alpha A_0(\phi) \in C\big([0,T], \dot{H}^1\big) \quad \textit{for all} \quad j \le m \quad \textit{and all} \quad |\alpha| \le M,$$

and $\left\|\partial_t^j \partial_x^\alpha A_0(\phi)\right\|_{L^\infty([0,T],\dot{H}^1)}$ is bounded by a continuous function of the norms $\left\|\partial_t^k \phi\right\|_{L^\infty([0,T],H^M)}$ for $k \le m+1$.

4 Elliptic estimates

Our object here is to prove Lemmas 1, 4 and 5.

4.1 Basic estimates

We first prove existence and uniqueness for the equation

$$\Delta u - |\phi|^2 u = -\Im(\phi f) \tag{75}$$

on $\mathbb{R}^4$.

Lemma 6. *Let $\phi \in \dot{H}^1$ and $f \in L^2$. Then the equation* (75) *has a unique (real-valued) solution $u \in \dot{H}^1$, and*

$$\|u\|_{\dot{H}^1} \le 2 \|f\|_{L^2}. \tag{76}$$

Proof. Recall that $\dot{H}^1$, as defined in section 1.4, is a Hilbert space with inner product $\int \nabla u \cdot \overline{\nabla v}$ (by Plancherel's theorem), and that $\dot{H}^1 \hookrightarrow L^4$. We denote by $\Re\dot{H}^1$ the corresponding real Hilbert space, with inner product $\int \nabla u \cdot \nabla v$.

By definition, $u \in \dot{H}^1$ solves (75) in the sense of distributions iff

$$\int_{\mathbb{R}^4} \big(\nabla u \cdot \nabla v + |\phi|^2 uv\big)\, dx = \int \Im(\phi f) v \, dx \tag{77}$$

for all $v \in \mathcal{S}$. Since $\mathcal{S}$ is dense in $\dot{H}^1$ and

$$\left| \int |\phi|^2 \, uv \, dx \right| \le \|\phi\|_{L^4}^2 \, \|u\|_{L^4} \, \|v\|_{L^4} \lesssim \|\phi\|_{\dot{H}^1}^2 \, \|u\|_{\dot{H}^1} \, \|v\|_{\dot{H}^1} \,, \tag{78}$$

$$\left| \int \Im(\phi f) v \, dx \right| \le \|\phi\|_{L^4} \, \|f\|_{L^2} \, \|v\|_{L^4} \lesssim \|\phi\|_{\dot{H}^1} \, \|v\|_{\dot{H}^1} \, \|f\|_{L^2} \,, \tag{79}$$

we conclude that u solves (75) iff (77) holds for all $v \in \dot{H}^1$. Taking $v = \overline{u}$ gives

$$\|u\|_{\dot{H}^1}^2 + \|\phi u\|_{L^2}^2 = \int \Im(\phi f) \overline{u} \, dx \le \|\phi u\|_{L^2} \, \|f\|_{L^2} \,,$$

and since $(a+b)^2 \le 2(a^2+b^2)$ for all $a, b \in \mathbb{R}$, we conclude that $N^2 \le 2N \|f\|_{L^2}$ where $N = \|u\|_{\dot{H}^1} + \|\phi u\|_{L^2} < \infty$. Therefore (76) holds, and uniqueness follows. Of course, u must be real, since if u solves (75), then $\Im u$ solves the same equation with $f = 0$, and therefore $\Im u = 0$ by what we just proved.

To prove existence, observe that the left hand side of (77) defines an inner product on $\Re \dot{H}^1$, and in view of (78), the corresponding norm is equivalent to the usual norm. Moreover, by (79), the right hand side of (77) is a bounded linear functional $F(v)$ on $\Re \dot{H}^1$. Existence therefore follows from the Riesz representation theorem. $\square$

Remark 3. As discussed in the introduction, our method can be modified to generalize the result of Cuccagna [1] for MKG on $\mathbb{R}^{1+3}$ to large data in H^s, $s > 3/4$. For this, we need the fact that (75) has a unique solution in $\dot{H}^1(\mathbb{R}^3)$ for $\phi \in H^{3/4}(\mathbb{R}^3)$ and $f \in H^{-1/4}(\mathbb{R}^3)$. Again we multiply the equation by $\overline{u}$ and integrate. Using Plancherel's theorem we get

$$\|u\|_{\dot{H}^1}^2 + \|\phi u\|_{L^2}^2 \le \|\phi u\|_{H^{1/4}} \, \|f\|_{H^{-1/4}} + \|\phi \overline{u}\|_{H^{1/4}} \, \|f\|_{H^{-1/4}}$$

and since

$$\|\phi u\|_{H^{1/4}} \lesssim \|\phi\|_{H^{3/4}} \, \|u\|_{\dot{H}^1} \,,$$

on $\mathbb{R}^3$, we get $\|u\|_{\dot{H}^1} \lesssim \|\phi\|_{H^{3/4}} \, \|f\|_{H^{-1/4}}$. It is also easy to show that the operator B_0 defined by (30) is bounded in L^2 for $\phi, A_j \in H^{3/4}(\mathbb{R}^3)$.

Next, we prove a difference estimate for (75).

Lemma 7. *Let $\phi, \psi \in \dot{H}^1$ and $f, g \in L^2$. Let $u, v \in \dot{H}^1$ be the solutions of*

$$\Delta u - |\phi|^2 \, u = -\Im(\phi f),$$
$$\Delta v - |\psi|^2 \, v = -\Im(\psi g).$$

Then

$$\|u - v\|_{\dot{H}^1} \lesssim \|\phi - \psi\|_{\dot{H}^1} + \|f - g\|_{L^2}$$

where the suppressed constant is a polynomial in $\|\phi\|_{\dot{H}^1}$, $\|\psi\|_{\dot{H}^1}$ and $\|g\|_{L^2}$.

Proof. Subtracting the equations gives

$$\Delta(u-v) - |\phi|^2 (u-v) = \big(|\phi|^2 - |\psi|^2\big)v - \Im\left[\phi(f-g)\right] - \Im\left[(\phi-\psi)g\right].$$

Then by a density argument as in the previous proof,

$$\begin{aligned}\int \Big(\nabla(u-v)\cdot\nabla(u-v) + |\phi|^2 (u-v)^2\Big)\,dx \qquad\qquad\qquad\qquad\\ = \int \Big(\big(|\psi|^2 - |\phi|^2\big)v + \Im\left[\phi(f-g)\right] + \Im\left[(\phi-\psi)g\right]\Big)(u-v)\,dx\\ \leq \|\phi-\psi\|_{L^4}\big(\|\phi\|_{L^4} + \|\psi\|_{L^4}\big)\|v\|_{L^4}\|u-v\|_{L^4}\\ + \|\phi\|_{L^4}\|f-g\|_{L^2}\|u-v\|_{L^4} + \|\phi-\psi\|_{L^4}\|g\|_{L^2}\|u-v\|_{L^4},\end{aligned}$$

giving the desired conclusion. □

We now consider the more general equation

$$\Delta u - |\phi|^2 u = f \tag{80}$$

Lemma 8. *Given $\phi \in \dot{H}^1$ and $f \in L^{4/3}$, the equation* (80) *has a unique solution $u \in \dot{H}^1$, and*

$$\|u\|_{\dot{H}^1} \leq C\,\|f\|_{L^{4/3}} \tag{81}$$

where C is independent of ϕ, f and u. Moreover, if

$$\Delta u - |\phi|^2 u = f, \tag{82}$$

$$\Delta v - |\psi|^2 v = g, \tag{83}$$

where $u, v, \phi, \psi \in \dot{H}^1$ and $f, g \in L^{4/3}$, then

$$\|u-v\|_{\dot{H}^1} \leq C\big(\|\phi\|_{\dot{H}^1} + \|\psi\|_{\dot{H}^1}\big)\|g\|_{L^{4/3}}\|\phi-\psi\|_{\dot{H}^1} + C\,\|f-g\|_{L^{4/3}},$$

with the same constant C as above.

Proof. Proceed as in the proof of Lemma 6, but with the right hand side of (77) replaced by $-\int vf\,dx$. Thus (79) is replaced by

$$\left|\int vf\,dx\right| \leq \|v\|_{L^4}\|f\|_{L^{4/3}} \lesssim \|v\|_{\dot{H}^1}\|f\|_{L^{4/3}}.$$

Existence then follows, and any $\dot{H}^1$ solution satisfies

$$\|u\|_{\dot{H}^1}^2 + \|\phi u\|_{L^2}^2 = -\int uf\,dx \leq C\,\|u\|_{\dot{H}^1}\|f\|_{L^{4/3}},$$

where C is independent of u, f and ϕ, and (81) follows.

Subtracting (82) from (81) gives

$$\big(\Delta - |\phi|^2\big)(u-v) = \big(|\phi|^2 - |\psi|^2\big)v + f - g,$$

and applying (81) gives the desired estimate. □

Next we prove a uniqueness result in space-time:

Lemma 9. *Suppose*

$$\phi \in C\big([0,T], \dot{H}^1\big) \quad \textit{and} \quad u \in L^2\big([0,T], \dot{H}^1\big),$$

and that u *solves*

$$\Delta u - |\phi|^2 u = 0 \quad \textit{on} \quad (0,T) \times \mathbb{R}^4$$

in the sense of distributions. Then $u = 0$.

Proof. Set $S_T = (0,T) \times \mathbb{R}^4$. For every test function $v(t,x)$ in $C_c^\infty(S_T)$,

$$\int \big\{ \nabla u \cdot \nabla v + |\phi|^2 uv \big\} \, dt \, dx = 0. \tag{84}$$

The left hand side is a bounded linear functional in v. In fact,

$$\left| \int \nabla u \cdot \nabla v \, dt \, dx \right| \le \|\nabla u\|_{L^2(S_T)} \|\nabla v\|_{L^2(S_T)} = \|u\|_{L_t^2(\dot{H}^1)} \|v\|_{L_t^2(\dot{H}^1)} \,,$$
$$\left| \int |\phi|^2 uv \, dt \, dx \right| \lesssim \|\phi\|^2_{L_t^\infty(\dot{H}^1)} \|u\|_{L_t^2(\dot{H}^1)} \|v\|_{L_t^2(\dot{H}^1)} \,.$$

Here we used Hölder's inequality and the embedding $\dot{H}^1 \hookrightarrow L^4$.

But $C_c^\infty(S_T)$ is dense in $L^2([0,T], \dot{H}^1)$, so it follows that (84) must hold for all $v \in L^2([0,T], \dot{H}^1)$. Taking $v = \overline{u}$ gives

$$\int \big\{ |\nabla u|^2 + |\phi|^2 |u|^2 \big\} \, dt \, dx = 0.$$

This implies $\nabla u = 0$, hence $u = 0$ ($\dot{H}^1$, as we have defined it, does not contain any nonzero constants). □

4.2 Higher regularity estimates

Suppose

$$\phi \in C\big([0,T], \dot{H}^1\big) \quad \text{and} \quad f \in C\big([0,T], L^{4/3}\big).$$

By Lemma 8, the equation

$$\Delta u - |\phi|^2 u = f \tag{85}$$

has a unique solution

$$u \in C\big([0,T], \dot{H}^1\big).$$

We shall prove the following higher regularity estimate.

Lemma 10. *Let* m, M *be non-negative integers. If*

$$\partial_t^j \partial_x^\alpha \phi \in C\big([0,T], \dot{H}^1\big) \quad \textit{and} \quad \partial_t^j \partial_x^\alpha f \in C\big([0,T], L^{4/3}\big)$$

for all $j \le m$ and $|\alpha| \le M$, then

$$\partial_t^j \partial_x^\alpha u \in C\big([0,T], \dot H^1\big)$$

for $j \le m$ and $|\alpha| \le M$, and

$$\begin{aligned}\left\|\partial_t^j \partial_x^\alpha u\right\|_{L_t^\infty(\dot H^1)} \lesssim \left\|\partial_t^j \partial_x^\alpha \phi\right\|_{L_t^\infty(\dot H^1)} \|\phi\|_{L_t^\infty(\dot H^1)} \|f\|_{L_t^\infty(L^{4/3})} \\ + \left\|\partial_t^j \partial_x^\alpha f\right\|_{L_t^\infty(L^{4/3})} + \eta_\alpha,\end{aligned}$$

where η_α is a lower order term which depends continuously on the norms

$$\left\|\partial_t^k \partial_x^\beta \phi\right\|_{L_t^\infty(\dot H^1)} \quad \text{and} \quad \left\|\partial_t^k \partial_x^\beta f\right\|_{L_t^\infty(L^{4/3})}$$

for all $k \le j$ and $|\beta| \le |\alpha|$ satisfying $k + |\beta| < j + |\alpha|$. Here all L^∞-norms are taken over $[0,T]$.

The proof is by induction on m and M. Denote by $P(m,M)$ the statement that the lemma holds for the pair (m,M). Since $P(0,0)$ is true by Lemma 8, it is enough, by an obvious induction, to prove

$$P(m,0) \implies P(m+1,0), \tag{86a}$$
$$P(0,M) \implies P(0,M+1), \tag{86b}$$
$$P(m+1,M), P(m,M+1) \implies P(m+1,M+1). \tag{86c}$$

The key to proving these implications is the following:

Lemma 11. *If, for some $0 \le \mu \le 4$,*

$$\partial_\mu \phi \in C\big([0,T], \dot H^1\big) \quad \text{and} \quad \partial_\mu f \in C\big([0,T], L^{4/3}\big),$$

then $\partial_\mu u \in C([0,T], \dot H^1)$ and

$$\|\partial_\mu u\|_{L_t^\infty(\dot H^1)} \lesssim \|f\|_{L_t^\infty(L^{4/3})} \|\phi\|_{L_t^\infty(\dot H^1)} \|\partial_\mu \phi\|_{L_t^\infty(\dot H^1)} + \|\partial_\mu f\|_{L_t^\infty(L^{4/3})},$$

where the L^∞-norms are taken over $[0,T]$.

Before proving Lemma 11, let us use it to prove (86).

Proof of (86a)**.** Apply ∂_t^m to both sides of (85). This gives

$$\big(\Delta - |\phi|^2\big)\partial_t^m u = \sum_{j+k+l=m,\; l<m} c_{jkl} (\partial_t^j \phi)(\overline{\partial_t^k \phi})(\partial_t^l u) + \partial_t^m f. \tag{87}$$

Denote the right hand side of this equation by F. By Lemma 11, if we can show that F and $\partial_t F$ belong to $C([0,T], L^{4/3})$, then it follows that

$$\partial_t^{m+1} u \in C([0,T], \dot H^1)$$

and

$$\left\|\partial_t^{m+1}u\right\|_{L_t^\infty(\dot H^1)} \lesssim \|F\|_{L_t^\infty(L^{4/3})}\|\phi\|_{L_t^\infty(\dot H^1)}\|\partial_t\phi\|_{L_t^\infty(\dot H^1)} + \|\partial_t F\|_{L_t^\infty(L^{4/3})}.$$

But by Hölder's inequality and the embedding $\dot H^1 \hookrightarrow L^4$,

$$\|F\|_{L_t^\infty(L^{4/3})} \lesssim \sum \left\|\partial_t^j\phi\right\|_{L_t^\infty(\dot H^1)}\left\|\partial_t^k\phi\right\|_{L_t^\infty(\dot H^1)}\left\|\partial_t^l u\right\|_{L_t^\infty(\dot H^1)} + \left\|\partial_t^m f\right\|_{L_t^\infty(L^{4/3})},$$

and using the hypothesis $P(m,0)$ to bound $\left\|\partial_t^l u\right\|_{L_t^\infty(\dot H^1)}$, we conclude that $\|F\|_{L_t^\infty(L^{4/3})}$ is bounded by a continuous function of the norms $\left\|\partial_t^j\phi\right\|_{L_t^\infty(\dot H^1)}$ and $\left\|\partial_t^j f\right\|_{L_t^\infty(L^{4/3})}$ for all $j \le m$.

Next apply ∂_t to F. Since $l < m$ in (87), we can again use the hypothesis $P(m,0)$ to bound $\left\|\partial_t^{l+1}u\right\|_{L_t^\infty(\dot H^1)}$. Thus, arguing as before,

$$\|\partial_t F\|_{L_t^\infty(L^{4/3})} \lesssim \left\|\partial_t^{m+1}\phi\right\|_{L_t^\infty(\dot H^1)}\|\phi\|_{L_t^\infty(\dot H^1)}\|f\|_{L_t^\infty(L^{4/3})} + \left\|\partial_t^{m+1}f\right\|_{L_t^\infty(L^{4/3})} + \text{l.o.t.},$$

where l.o.t. stands for a lower order term depending continuously on the norms $\left\|\partial_t^j\phi\right\|_{L_t^\infty(\dot H^1)}$ and $\left\|\partial_t^j f\right\|_{L_t^\infty(L^{4/3})}$ for all $j \le m$.

Thus $P(m+1,0)$ holds, completing the proof of (86a). The proof of (86b) is quite similar and is omitted.

Proof of (86c)**.** Let $|\alpha| = M+1$. Applying $\partial_t^m\partial_x^\alpha$ to both sides of (85) gives

$$\left(\Delta - |\phi|^2\right)\partial_t^m\partial_x^\alpha u = \sum c_{\beta\gamma\delta jkl}(\partial_t^j\partial_x^\beta\phi)(\overline{\partial_t^k\partial_x^\gamma\phi})(\partial_t^l\partial_x^\delta u) + \partial_t^m\partial_x^\alpha f, \tag{88}$$

where the sum is over all non-negative integers j,k,l and multi-indices β,γ,δ such that

$$\beta+\gamma+\delta = \alpha, \quad j+k+l = m, \quad l+|\delta| \le m+M.$$

Let F be the right hand side of (88). If F and $\partial_t F$ belong to $C([0,T],L^{4/3})$, then by Lemma 11,

$$\partial_t^{m+1}\partial_x^\alpha u \in C([0,T],\dot H^1)$$

and

$$\left\|\partial_t^{m+1}\partial_x^\alpha u\right\|_{L_t^\infty(\dot H^1)} \lesssim \|F\|_{L_t^\infty(L^{4/3})}\|\phi\|_{L_t^\infty(\dot H^1)}\|\partial_t\phi\|_{L_t^\infty(\dot H^1)} + \|\partial_t F\|_{L_t^\infty(L^{4/3})}.$$

Since F and $\partial_t F$ are estimated just as in the proof of (86a), we will not go into details. The key point is that since $l \le m$, $|\delta| \le M+1$ and $l+|\delta| \le m+M$ in (88), the hypotheses $P(m+1,M)$ and $P(m,M+1)$ allow us to bound $\left\|\partial_t^l\partial_x^\delta u\right\|_{L_t^\infty(\dot H^1)}$ and $\left\|\partial_t^{l+1}\partial_x^\delta u\right\|_{L_t^\infty(\dot H^1)}$.

Proof of Lemma 11. Under the hypotheses of the lemma,

$$\partial_\mu u \in L^2\big([0,T], \dot{H}^1\big). \tag{89}$$

Before proving this, let us show that it implies the conclusion of the lemma.

Indeed, u solves (85) in the sense of distributions on $S_T = (0,T) \times \mathbb{R}^4$, and if we apply ∂_μ to both sides, it follows that

$$\big(\Delta - |\phi|^2\big)\partial_\mu u = \partial_\mu\big(|\phi|^2\big)u + \partial_\mu f. \tag{90}$$

in the sense of distributions on S_T. (The use of the product rule for derivatives is easily justified in view of (89).) Denote by F the right hand side of the last equation. Then

$$F = 2\Re\big(\phi\overline{\partial_\mu\phi}\big)u + \partial_\mu f,$$

and so

$$\|F\|_{L_t^\infty(L^{4/3})} \lesssim \|\phi\|_{L_t^\infty(\dot{H}^1)} \|\partial_\mu\phi\|_{L_t^\infty(\dot{H}^1)} \|u\|_{L_t^\infty(\dot{H}^1)} + \|\partial_\mu f\|_{L_t^\infty(L^{4/3})}.$$

It then follows by Lemma 8 that the equation

$$\big(\Delta - |\phi|^2\big)v = F \tag{91}$$

has a solution $v \in C([0,T], \dot{H}^1)$, and

$$\|v\|_{L_t^\infty(\dot{H}^1)} \lesssim \|f\|_{L_t^\infty(L^{4/3})} \|\phi\|_{L_t^\infty(\dot{H}^1)} \|\partial_\mu\phi\|_{L_t^\infty(\dot{H}^1)} + \|\partial_\mu f\|_{L_t^\infty(L^{4/3})}.$$

Subtracting (91) from (90) gives

$$\big(\Delta - |\phi|^2\big)(\partial_\mu u - v) = 0.$$

Thus $\partial_\mu u = v$ by Lemma 9, proving the conclusion of Lemma 11.

It remains to prove (89). For technical reasons, we fix $0 < t_0 < T$ and prove (89) with the interval $[0,T]$ replaced by $[0,t_0]$. A similar argument works for the interval $[t_0,T]$, giving the statement in the entire interval $[0,T]$.

We shall require the following facts about the difference quotients

$$\Delta_h^j u(t,x) = \frac{u(t,x+he_j) - u(t,x)}{h} \quad \text{and} \quad \Delta_h^0 u(t,x) = \frac{u(t+h,x) - u(t,x)}{h},$$

where $e_1, \dots, e_4$ are the standard basis vectors of $\mathbb{R}^4$.

Lemma 12. *If f belongs to $C([0,T], L^{4/3})$ and the distributional derivative $\partial_\mu f$ belongs to $C([0,T], L^{4/3})$, for some $0 \le \mu \le 4$, then*

$$\|\Delta_h^\mu f\|_{L^\infty([0,t_0],L^{4/3})} = O(1) \quad \text{as} \quad h \to 0^+.$$

Moreover, the same conclusion holds with f replaced by ϕ and $L^{4/3}$ by $\dot{H}^1$.

Proof. We have

$$\Delta_h^0 f(t) = \frac{1}{h}\int_0^h \partial_t f(t+s)\,ds \qquad (L^{4/3}\text{-valued integral})$$

whence

$$\left\|\Delta_h^0 f(t)\right\|_{L^{4/3}} \le \sup_{0\le s\le h} \|\partial_t f(t+s)\|_{L^{4/3}}.$$

The same proof works for $\dot H^1$. If $1 \le j \le 4$, then (discarding the time variable)

$$\Delta_h^j f(x) = \frac{1}{h}\int_0^h \partial_j f(x+se_j)\,ds.$$

Thus, by Minkowski's integral inequality and the translation invariance of the norm,

$$\left\|\Delta_h^j f\right\|_{L^{4/3}} \le \|\partial_j f\|_{L^{4/3}}.$$

This is certainly valid for smooth f, and hence in general by using an approximation of the identity. For $\dot H^1$ we write

$$\left\|\Delta_h^j \phi\right\|_{\dot H^1}^2 = \int |\xi|^2 \left|\frac{e^{ih\xi_j}-1}{h}\right|^2 \left|\widehat{\phi}\right|^2 d\xi$$

and note that $\left|e^{ih\xi_j}-1\right| \le h\,|\xi_j|$. It follows that $\left\|\Delta_h^j \phi\right\|_{\dot H^1} \le \|\partial_j \phi\|_{\dot H^1}$. $\square$

We are now ready to prove (89). By the difference estimate in Lemma 8,

$$\begin{aligned}\|\Delta_h^\mu u\|_{L^\infty([0,t_0],\dot H^1)} \lesssim \|f\|_{L^\infty([0,T],L^{4/3})} \|\phi\|_{L^\infty([0,T],\dot H^1)} \|\Delta_h^\mu \phi\|_{L^\infty([0,t_0],\dot H^1)} \\ + \|\Delta_h^\mu f\|_{L^\infty([0,t_0],L^{4/3})},\end{aligned}$$

for all $0 < h < T - t_0$. In view of Lemma 12, the right hand side is $O(1)$ as $h \to 0^+$. Applying Hölder's inequality in time then gives

$$\|\Delta_h^\mu u\|_{L^2([0,t_0],\dot H^1)} = O(1) \quad \text{as} \quad h \to 0^+,$$

so by weak compactness, there is a sequence $h_j \to 0$ such that $\Delta_{h_j}^\mu u$ converges weakly in $L^2([0,t_0],\dot H^1)$ to some limit v as $j \to \infty$. But this implies that $\Delta_{h_j}^\mu u \to v$ also in the sense of distributions on $(0,t_0)\times\mathbb{R}^4$. On the other hand, we know that $\Delta_h^\mu u \to \partial_\mu u$ in the distributional sense as $h \to 0$, and so we conclude that $\partial_\mu u = v$. This proves (89) on the interval $[0,t_0]$.

4.3 Proofs of Lemmas 1, 4 and 5

First, Lemma 1 is an immediate corollary of Lemmas 6 and 7.

Secondly, to prove Lemma 4, we apply Lemma 10 with $m = 0$, $M = k$ and

$$f = -\Im\big(\phi\overline{\partial_t\phi}\big).$$

Since $\Lambda^k\phi \in \mathcal{X}_2 = \mathcal{H}^{s,\theta}$ by hypothesis, we have, in view of (17),

$$\partial_x^\alpha \phi \in C\big([0,T], \dot{H}^1\big) \quad \text{and} \quad \partial_t \partial_x^\alpha \phi \in C\big([0,T], L^2\big) \quad \text{for} \quad |\alpha| \le k. \tag{92}$$

Thus, it suffices to check that $\partial_x^\alpha f \in C([0,T], L^{4/3})$ for $|\alpha| \le k$, and

$$\|\partial_x^\alpha f\|_{L_t^\infty(L^{4/3})} \lesssim \|\phi\|_{\mathcal{X}_2} \|\partial_x^\alpha \phi\|_{\mathcal{X}_2} + \eta\Big(\sum_{|\beta|<|\alpha|} \big\|\partial_x^\beta \phi\big\|_{\mathcal{X}_2}\Big), \tag{93}$$

where η is continouous. If $\alpha = 0$, we have, using Hölder's inequality and the embedding $\dot{H}^1 \hookrightarrow L^4$,

$$\|f\|_{L^{4/3}} \le \|\phi\|_{L^4} \|\partial_t \phi\|_{L^2} \lesssim \|\phi\|_{\dot{H}^1} \|\partial_t\phi\|_{L^2}$$

uniformly in $0 \le t \le T$, and by (17),

$$\|\phi\|_{\dot{H}^1}, \|\partial_t \phi\|_{L^2} \lesssim \|\phi\|_{\mathcal{X}_2},$$

giving (93) for $\alpha = 0$. When $\alpha \neq 0$ one can apply the product rule and estimate each term as above. We leave the details to the interested reader.

Finally, Lemma 5 is also proved by an application of Lemma 10. We are given non-negative integers m, M such that

$$\partial_t^j \partial_x^\alpha \phi \in C\big([0,T], L^2\big) \quad \text{for all} \quad j \le m+1 \quad \text{and all} \quad |\alpha| \le M+1.$$

Again we set

$$f = -\Im\big(\phi \overline{\partial_t \phi}\big).$$

By Lemma 10 it suffices to check that

$$\partial_t^j \partial_x^\alpha f \in C\big([0,T], L^{4/3}\big) \quad \text{for all} \quad j \le m \quad \text{and all} \quad |\alpha| \le M.$$

Again, one simply applies the product rule for derivatives and estimates each term as in the proof of (93).

Appendix

Here we prove (54). First,

$$\big\|\Lambda^{s-1}(-\Delta)^{-1}(uv)\big\|_{L_t^p(L_x^q)} \lesssim \big\|(-\Delta)^{-1}(uv)\big\|_{L_t^p(L_x^q)} + \big\|(-\Delta)^{\frac{s-3}{2}}(uv)\big\|_{L_t^p(L_x^q)}$$

by Lemma 3, so it suffices to show that the two terms on the right hand side are both $\lesssim \|u\|_{H^{s,\theta}} \|v\|_{H^{s-1,\theta}}$. To this end, we apply the following theorem (stated here for $\mathbb{R}^{1+4}$ only) of Klainerman-Tataru [8]:

Theorem. *Let $1 \leq p \leq \infty$, $1 \leq q < \infty$ and set $\gamma = 2 - \frac{1}{2p} - \frac{2}{q}$. Assume that*

$$\frac{1}{p} \leq \frac{3}{2}\left(1 - \frac{1}{q}\right), \tag{94}$$

$$0 < \sigma < 4 - \frac{2}{p} - \frac{4}{q}, \tag{95}$$

$$s_1, s_2 < \gamma, \tag{96}$$

$$s_1 + s_2 + \sigma = 2\gamma. \tag{97}$$

Then

$$\left\| (-\Delta)^{-\sigma/2}(uv) \right\|_{L_t^p(L_x^q)} \lesssim \|u\|_{H^{s_1,\theta}} \|v\|_{H^{s_2,\theta}} ,$$

where $\theta > \frac{1}{2}$.

Remarks. (1) When $s_1 = s_2$ this follows from Theorem 4 in [8] (see also [7, Principle 3.2]). In [8], however, the estimate was stated using the space-time fractional derivative operator $(-\Delta_{t,x})^{-\sigma/2}$. Nevertheless, an inspection of their proof shows that it works equally well for $(-\Delta)^{-\sigma/2}$ (see [11, Chapter 2]). In our statement of the theorem we have also included the end-point case due to Keel and Tao [3], although we do not use this.

(2) The asymmetric case $s_1 \neq s_2$ is derived as in the proof of Theorem 5 in [8]. (The statement of that theorem contains the condition (in our notation) $\sigma \leq \gamma$, but an inspection of the proof shows that this is superfluous.) Let us just give a heuristic explanation of why the asymmetric case essentially reduces to the symmetric situation. Rewrite the estimate as follows:

$$\left\| D^{-\sigma}(D^{-s_1}u \cdot D^{-s_2}v) \right\|_{L_t^p(L_x^q)} \lesssim \|u\|_{H^{0,\theta}} \|v\|_{H^{0,\theta}} ,$$

where $D^\alpha = (-\Delta)^{\alpha/2}$. Denote by ξ and η the Fourier frequencies of u and v corresponding to the spatial variable x. Then the frequency of the product uv is $\xi + \eta$, and in Fourier space, $D^{-\sigma}(D^{-s_1}u \cdot D^{-s_2}v)$ is a weighted convolution, with weights

$$\frac{1}{|\xi+\eta|^\sigma |\xi|^{s_1} |\eta|^{s_2}}$$

The idea is that the weights can be redistributed so as to get equal powers of $|\xi|$ and $|\eta|$. This is obviously possible if the frequencies of u and v are comparable. If, on the other hand, $|\xi| \gg |\eta|$, say, then $|\xi + \eta| \sim |\xi|$, and so

$$\frac{1}{|\xi+\eta|^\sigma |\xi|^{s_1} |\eta|^{s_2}} \lesssim \frac{1}{|\xi|^\gamma |\eta|^\gamma}$$

provided $s_1, s_2 \leq \gamma$ (recall that $s_1 + s_2 + \sigma = 2\gamma$). Thus we are in the case $\sigma = 0$ and $s_1 = s_2 = \gamma$, which by Hölder's inequality is reduced to a linear Strichartz estimate.

Now let p and q be defined as in section 2. Using the definition of p in (50), we see that (94) is equivalent to $1/q \le (2/3)(\theta + 2\varepsilon)$, and the latter evidently holds, since $1/q \le 1/4$ by (51), (49a) and the assumption $s < 2$. Thus (94) holds.

Next we have to check that (95) holds with $\sigma = 2$ (then it also holds with $\sigma = 3 - s$, of course), but using the definition of p in (50), we find that (95) is equivalent to

$$\frac{4}{q} < 2\theta - 1 + 4\varepsilon,$$

which is true by (51).

Now set $s_1 = 1 + \delta$ and $s_2 = \delta$, where we have defined

$$\delta = \frac{3-\sigma}{2} - \frac{1}{2p} - \frac{2}{q}. \tag{98}$$

With this choice, (97) clearly holds. Note that (96) holds provided

$$\delta < 1 - \frac{1}{2p} - \frac{2}{q}.$$

It suffices to check this when $\sigma = 3 - s$ (then it also holds for $\sigma = 2$), but in this case it is obvious since $s < 2$ and, from (98),

$$\delta = \frac{s}{2} - \frac{1}{2p} - \frac{2}{q}. \tag{99}$$

It remains to check that $s_1 \le s$ and $s_2 \le s-1$. This is equivalent to $\delta \le s-1$, and again we only have to check this for $\sigma = 3 - s$, in which case it reduces to, by (99),

$$\frac{s}{2} - \frac{1}{2p} - \frac{2}{q} \le s - 1.$$

In fact,

$$\frac{s}{2} - \frac{1}{2p} < s - 1,$$

for by (50), this is equivalent to $1/4 + \theta/2 + \varepsilon < s/2$, which holds by (49b).

References

[1] S. Cuccagna, *On the local existence for the Maxwell-Klein-Gordon system in* R^{3+1}, Comm. PDE **24** (1999), no. 5-6, 851–867

[2] D. Eardley and V. Moncrief, *The global existence of Yang-Mills-Higgs fields in* R^{3+1}, Comm. Math. Phys. **83** (1982), 171–212

[3] M. Keel and T. Tao, *Endpoint Strichartz estimates*, Amer. J. Math. **120** (1998), no. 5, 955–980

[4] S. Klainerman and M. Machedon, *On the Maxwell-Klein-Gordon equation with finite energy*, Duke Math. J. **74** (1994), 19–44

[5] S. Klainerman and M. Machedon, *Remark on Strichartz type inequalities*, Int. Math. Res. Not., no. 5 (1996), 201–220

[6] S. Klainerman and M. Machedon, *On the optimal local regularity for gauge field theories*, Differential and Integral Equations **10** (1997), 1019–1030

[7] S. Klainerman and S. Selberg, *Bilinear estimates and applications to nonlinear wave equations*, to appear in Comm. in Contemporary Math.

[8] S. Klainerman and D. Tataru, *On the optimal local regularity for Yang-Mills equations in* $\mathbb{R}^{4+1}$, J. Amer. Math. Soc. **12** (1999), 93–116

[9] G. Ponce and T. Sideris, *Local regularity of nonlinear wave equations in three space dimensions*, Comm. PDE **18** (1993), 169–177

[10] H. Lindblad, *Counterexamples to local existence for semilinear wave equations*, Amer. J. Math. **118** (1996), 1–16

[11] S. Selberg, *Multilinear space-time estimates and applications to local existence theory for nonlinear wave equations*, Ph.D. Thesis, Princeton University 1999

[12] S. Selberg, *On an estimate for the wave equation and applications to nonlinear problems*, Differential and Integral Equations **2** (2002), 213–236

[13] C. D. Sogge, "Lectures on nonlinear wave equations", Monographs in Analysis, II. International Press, 1995

[14] E. Stein, "Singular Integrals and Differentiability Properties of Functions", Princeton University Press, 1970

[15] T. Tao, *Global regularity of wave maps I. Small critical Sobolev norm in high dimension*, IMRN **7** (2001), 299–328

[16] T. Tao, *Global regularity of wave maps II. Small energy in two dimensions*, to appear, Comm. Math. Phys.

[17] T. Tao, *Local well-posedness of the Yang-Mills equation in the temporal gauge below the energy norm*, preprint